Counting Primitive Partial Words*

F. Blanchet-Sadri[1] Mihai Cucuringu[2]

January 18, 2011

Abstract

A word is primitive if it is not a power of another word. The number of primitive words of a fixed length over an alphabet of a fixed size is well known and relates to the Möbius function. In this paper, we investigate the number of primitive partial words which are strings that may contain "do not know" symbols.

Keywords: Combinatorics on words; Words; Partial words; Primitive words; Primitive partial words; Möbius function; Periods; Exact periods.

1 Introduction

Primitive words, or strings over a finite alphabet that cannot be written as a power of another string, play an important role in numerous research areas including formal language theory [16, 17], coding theory [4, 26], and combinatorics on words [14, 20, 21, 22, 23]. *Partial words* (or *pwords*) are strings that may contain a number of "do not know" symbols also called "holes" (words, or *full words*, are partial words without holes). *Primitive partial words* were defined in [5]. A partial word u is primitive if there exists no word v such that $u \subset v^i$ with $i \geq 2$ (the concept of containment, denoted by $\subset$, is discussed in Section 2). Testing whether or not a partial word is primitive can be done in linear time in the length of the word [6]. This result, which extends a result on words [15], found a nice application in [9]. There, Blanchet-Sadri and Chriscoe extend to partial words with one hole the well known result of Guibas and Odlyzko [19] which states that the sets of periods of words are independent of the alphabet size. Other recent results on partial words appear in [1, 2, 3, 7, 8, 10, 12, 13, 25].

The number of primitive words of a fixed length over an alphabet of a fixed size is well known and relates to the Möbius function [20]. In this paper, we investigate the number of primitive partial words. In Section 2, we discuss the well-known formula for the number of primitive full words of length n over an alphabet of size k, and start counting primitive partial words by considering the case of prime length. Section 3 contains several definitions and some important general properties on periods and exact periods of partial words that are useful for later sections. We recall in particular Fine and Wilf's theorem in the framework of partial words. In Section 4, we present our first counting method which consists in first considering all nonprimitive pwords with h holes obtained by replacing h positions in nonprimitive full words with $\diamond$'s (representing holes), and then

*This material is based upon work supported by the National Science Foundation under Grant No. DMS–0452020. We thank the referee of a preliminary version of this paper for his/her very valuable comments and suggestions.

[1]Department of Computer Science, University of North Carolina, P.O. Box 26170, Greensboro, North Carolina 27402–6170, USA blanchet@uncg.edu.

[2]Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, New Jersey 08544-1000, USA.

subtracting the pwords that have been doubly counted. There, we express in particular the number of primitive partial words with one or two holes of length n over a k-size alphabet in terms of the number of such full words. Finally, Section 5 discusses our second method. We count nonprimitive partial words of length n with h holes over a k-size alphabet through a constructive method that refines the counting done in the previous sections.

We end this section by reviewing basic concepts on words and partial words. Let A be a nonempty finite set, or an *alphabet*. A *string* or *word* u over A is a finite concatenation of symbols from A. The number of symbols in u, or *length* of u, is denoted by $|u|$. We assume that, for every word, the first letter is at position 0. For any word u, $u[i..j)$ is the *subword* or *factor* of u that starts at position i and ends at position $j-1$ (it is called *proper* if $0 \leq i < j \leq |u|$ and $(0 < i$ or $j < |u|)$). In particular, $u[0..j)$ is the *prefix* of u that ends at position $j-1$ and $u[i..|u|)$ is the *suffix* of u that begins at position i. The subword $u[i..j)$ is the empty word if $i \geq j$ (the empty word is denoted by ε). The set of all words over A of finite length (greater than or equal to 0) is denoted by A^*. It is a monoid under the associative operation of concatenation or product of words (ε serves as identity) and is referred to as the *free monoid* generated by A. Similarly, the set of all nonempty words over A is denoted by A^+. It is a semigroup under the operation of concatenation of words and is referred to as the *free semigroup* generated by A.

A word of length n over A can be defined by a total function $u : \{0, \ldots, n-1\} \to A$ and is usually represented as $u = a_0 a_1 \ldots a_{n-1}$ with $a_i \in A$. A partial word (or pword) u of length n over A is a partial function $u : \{0, \ldots, n-1\} \to A$. For $0 \leq i < n$, if $u(i)$ is defined, then we say that i belongs to the *domain* of u (denoted by $i \in D(u)$), otherwise we say that i belongs to the *set of holes* of u (denoted by $i \in H(u)$). A word over A is a partial word over A with an empty set of holes (we will sometimes refer to words as *full* words). The length of u is denoted by $|u|$.

If u is a partial word of length n over A, then the *companion* of u, denoted by $u_\diamond$, is the total function $u_\diamond : \{0, \ldots, n-1\} \to A \cup \{\diamond\}$ defined by

$$u_\diamond(i) = \begin{cases} u(i) & \text{if } i \in D(u) \\ \diamond & \text{otherwise} \end{cases}$$

The bijectivity of the map $u \mapsto u_\diamond$ allows us to define for partial words concepts such as concatenation in a trivial way. The symbol $\diamond$ is viewed as a "do not know" symbol. The word $u_\diamond = \diamond ba \diamond abb$ is the companion of the partial word u of length 7 where $D(u) = \{1, 2, 4, 5, 6\}$ and $H(u) = \{0, 3\}$. In the sequel, for convenience, we will consider a partial word over A as a word over the enlarged alphabet $A \cup \{\diamond\}$, where the additional symbol $\diamond$ plays a special role. Thus, we say for instance "the partial word $\diamond ba \diamond abb$" instead of "the partial word with companion $\diamond ba \diamond abb$".

2 Primitive partial words

For a word u, the powers of u are defined inductively by $u^0 = \varepsilon$ and, for any $i \geq 1$, $u^i = uu^{i-1}$. The bijectivity of the map $u \mapsto u_\diamond$ allows us to define for partial words powers in a trivial way, that is, $(u^i)_\diamond = (u_\diamond)^i$. A word u is *primitive* if there exists no word v such that $u = v^i$ with $i \geq 2$. If u is a nonempty word, then there exists a unique primitive word v and a unique positive integer i such that $u = v^i$.

We can extend the notion of a word being primitive to a partial word being primitive as follows. First, if u and v are two partial words "of equal length", then u is said to be contained in v, denoted by $u \subset v$, if all elements in $D(u)$ are in $D(v)$ and $u(i) = v(i)$ for all $i \in D(u)$. Now, a partial word u is *primitive* if there exists no word v such that $u \subset v^i$ with $i \geq 2$. Note that if v is primitive and $v \subset u$, then u is primitive as well. If u is a nonempty partial word, then there exists a primitive word v and a positive integer i such that $u \subset v^i$. Uniqueness does not hold for partial words ($u = \diamond a$ serves as a counterexample ($u \subset a^2$ and $u \subset ba$ for distinct letters a, b)).

Denote by $P_{h,k}(n)$ (respectively, $N_{h,k}(n)$) the number of primitive (respectively, nonprimitive) partial words with h holes of length n over a k-size alphabet A. Also, denote by $\mathcal{P}_{h,k}(n)$ (respectively, $\mathcal{N}_{h,k}(n)$) the set of primitive (respectively, nonprimitive) partial words with h holes of length n over A. Let $T_{h,k}(n)$ denote the total number of partial words of length n with h holes over A, and $\mathcal{T}_{h,k}(n)$ the set of all such partial words. It holds true that

$$P_{h,k}(n) + N_{h,k}(n) = T_{h,k}(n)$$

and it is easy to see that

$$T_{h,k}(n) = \binom{n}{h} k^{n-h}$$

We first count primitive full words. Since there are exactly k^n words of length n over A and every nonempty word w has a unique primitive root v for which $w = v^{n/d}$ for some divisor d of n, the following relation holds:

$$k^n = \sum_{d|n} P_{0,k}(d)$$

Using the Möbius inversion formula, we obtain the following well-known expression for $P_{0,k}(n)$ (see, e.g., [20, 24, 27]):

$$P_{0,k}(n) = \sum_{d|n} \mu(d) k^{n/d}$$

where the Möbius function, denoted by μ, is defined as

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^i & \text{if } n \text{ is a product of } i \text{ distinct primes} \\ 0 & \text{if } n \text{ is divisible by the square of a prime} \end{cases}$$

We now count primitive partial words of prime length p. If w is a nonprimitive pword with h holes of length p, then $w \subset a^p$ for some letter $a \in A$. There are k choices for a and $\binom{p}{h}$ for the holes. Thus

$$N_{h,k}(p) = \binom{p}{h} k$$

$$P_{h,k}(p) = T_{h,k}(p) - N_{h,k}(p) = \binom{p}{h} (k^{p-h} - k)$$

The tables below contain some numerical values for alphabets of sizes $k = 2$ and $k = 3$ that will be useful in the sequel (prime numbers n are underlined). These tables were obtained by having a computer generate all possible partial words with zero, one, two or three holes, and count the number of primitive and nonprimitive such words.

n	$T_{0,2}(n)$	$P_{0,2}(n)$	$N_{0,2}(n)$	$T_{1,2}(n)$	$P_{1,2}(n)$	$N_{1,2}(n)$
1	2	2	0	1	1	0
2	4	2	2	4	0	4
3	8	6	2	12	6	6
4	16	12	4	32	16	16
5	32	30	2	80	70	10
6	64	54	10	192	132	60
7	128	126	2	448	434	14
8	256	240	16	1024	896	128
9	512	504	8	2304	2232	72
10	1024	990	34	5120	4780	340
11	2048	2046	2	11264	11242	22
12	4096	4020	76	24576	23664	912
13	8192	8190	2	53248	53222	26
14	16384	16254	130	114688	112868	1820
15	32768	32730	38	245760	245190	570
16	65536	65280	256	524288	520192	4096
17	131072	131070	2	1114112	1114078	34
18	262144	261576	568	2359296	2349072	10224
19	524288	524286	2	4980736	4980698	38
20	1048576	1047540	1036	10485760	10465040	20720

n	$T_{2,2}(n)$	$P_{2,2}(n)$	$N_{2,2}(n)$	$T_{3,2}(n)$	$P_{3,2}(n)$	$N_{3,2}(n)$
1	0	0	0	0	0	0
2	1	0	1	0	0	0
3	6	0	6	1	0	1
4	24	4	20	8	0	8
5	80	60	20	40	20	20
6	240	102	138	160	24	136
7	672	630	42	560	490	70
8	1792	1376	416	1792	1088	704
9	4608	4320	288	5376	4716	660
10	11520	10070	1450	15360	11920	3440
11	28160	28050	110	42240	41910	330
12	67584	62760	4824	112640	97920	14720
13	159744	159588	156	292864	292292	572
14	372736	361354	11382	745472	703528	41944
15	860160	856170	3990	1863680	1846470	17210
16	1966080	1936384	29696	4587520	4458496	129024
17	4456448	4456176	272	11141120	11139760	1360
18	10027008	9942408	84600	26738688	26312400	426288
19	22413312	22412970	342	63504384	63502446	1938
20	49807360	49615640	191720	149422080	148333200	1088880

n	$T_{0,3}(n)$	$P_{0,3}(n)$	$N_{0,3}(n)$	$T_{1,3}(n)$	$P_{1,3}(n)$	$N_{1,3}(n)$
1	3	3	0	1	1	0
2	9	6	3	6	0	6
3	27	24	3	27	18	9
4	81	72	9	108	72	36
5	243	240	3	405	390	15
6	729	696	33	1458	1260	198
7	2187	2184	3	5103	5082	21
8	6561	6480	81	17496	16848	648
9	19683	19656	27	59049	58806	243
10	59049	58800	249	196830	194340	2490
11	177147	177144	3	649539	649506	33
12	531441	530640	801	2125764	2116152	9612
13	1594323	1594320	3	6908733	6908694	39
14	4782969	4780776	2193	22320522	22289820	30702
15	14348907	14348640	267	71744535	71740530	4005
16	43046721	43040160	6561	229582512	229477536	104976
17	129140163	129140160	3	731794257	731794206	51
18	387420489	387400104	20385	2324522934	2324156004	366930
19	1162261467	1162261464	3	7360989291	7360989234	57
20	3486784401	3486725280	59121	23245229340	23244046920	1182420

n	$T_{2,3}(n)$	$P_{2,3}(n)$	$N_{2,3}(n)$	$T_{3,3}(n)$	$P_{3,3}(n)$	$N_{3,3}(n)$
1	0	0	0	0	0	0
2	1	0	1	0	0	0
3	9	0	9	1	0	1
4	54	12	42	12	0	12
5	270	240	30	90	60	30
6	1215	774	441	540	144	396
7	5103	5040	63	2835	2730	105
8	20412	18360	2052	13608	10368	3240
9	78732	77760	972	61236	59022	2214
10	295245	284850	10395	262440	239040	23400
11	1082565	1082400	165	1082565	1082070	495
12	3897234	3847284	49950	4330260	4183632	146628
13	13817466	13817232	234	16888014	16887156	858
14	48361131	48171774	189357	64481508	63805728	675780

3 Periodicity

In this section, we discuss the concepts of *period* and *exact period*.

3.1 Periods

A *period* of a partial word w over A is a positive integer p such that $w(i) = w(j)$ whenever $i, j \in D(w)$ and $i \equiv j \bmod p$. In such a case, we call w *p-periodic*. We will denote the set of periods of w by $\mathcal{P}(w)$. The fundamental periodicity result of Fine and Wilf can be stated as follows.

Theorem 1 ([18])**.** *If a word w is p-periodic and q-periodic and $|w| \geq p + q - \gcd(p, q)$, then w is $\gcd(p, q)$-periodic.*

The bound $L(0, p, q) = p+q-\gcd(p, q)$ turns out to be optimal, since, for example, *abaababaaba* has periods 5 and 8, has length $11 = 5 + 8 - \gcd(5, 8) - 1$, but does not have period 1.

Berstel and Boasson proved a variant of Theorem 1 for partial words with one hole.

Theorem 2 ([1])**.** *If a partial word w with one hole is p-periodic and q-periodic and $|w| \geq p + q$, then w is $\gcd(p, q)$-periodic.*

The bound $L(1, p, q) = p + q$ turns out to be optimal since, for example, $aaaabaaaa\diamond aa$ has one hole, is 5-periodic and 8-periodic, has length $12 = 5 + 8 - 1$, but is not 1-periodic. Shur and Gamzova proved the following result.

Theorem 3 ([25])**.** *Let $p < q$ be positive integers.*

1. *If a partial word w with two holes is p-periodic and q-periodic and $|w| \geq 2p + q - \gcd(p, q)$, then w is $\gcd(p, q)$-periodic.*
2. *If a partial word w with h holes is p-periodic and q-periodic and $|w| \geq (h+1)p+q-\gcd(p, q)$, then w is $\gcd(p, q)$-periodic.*

Referring to Theorem 3, we denote the bound $2p + q - \gcd(p, q)$ by $L(2, p, q)$ which is optimal since, for example, $aba\diamond\diamond abaababaaba$ has two holes, is 5-periodic and 8-periodic, has length $16 = 2(5)+8-\gcd(5, 8)-1$, but is not 1-periodic. The following result gives an optimal bound $L(h, p, q)$ for h holes when q is large enough.

Theorem 4 ([7])**.** *Let p, q be positive integers satisfying $q > x(p, h)$ where*

$$x(p,h) = \begin{cases} p\left(\frac{h}{2}\right) & \text{if } h \text{ is even} \\ p\left(\frac{h+1}{2}\right) & \text{if } h \text{ is odd} \end{cases}$$

If a partial word w with h holes is p-periodic and q-periodic and $|w| \geq L(h, p, q)$, then w is $\gcd(p, q)$-periodic where

$$L(h,p,q) = \begin{cases} p\left(\frac{h+2}{2}\right) + q - \gcd(p,q) & \text{if } h \text{ is even} \\ p\left(\frac{h+1}{2}\right) + q & \text{if } h \text{ is odd} \end{cases}$$

In the case of three holes, $L(3, p, q) = 2p + q$ if $q > 2p$.

For other extensions of Fine and Wilf's periodicity result in the context of partial words, we refer the reader to [3, 12].

3.2 Exact periods

Consider a partial word $w = a_0 \ldots a_{n-1}$ where $a_i \in A \cup \{\diamond\}$. We call w an $\frac{n}{d}$-*repeat* if w is d-periodic and d is a divisor of n. We call the $\diamond$ in position $i \in H(w)$ *free with respect to period d* if whenever $j \in D(w)$, we have $j \not\equiv i \bmod d$. For example, if $w = aba\diamond aba\diamond$ then the $\diamond$'s in positions 3 and 7 are free with respect to period $d = 4$ but not free with respect to period $d = 2$:

$$\begin{array}{cccc} a & b & a & \diamond \\ a & b & a & \diamond \end{array} \qquad \begin{array}{cc} a & b \\ a & \diamond \\ a & b \\ a & \diamond \end{array}$$

We denote by $\mathcal{D}(n)$ the set of divisors of n distinct from n, by $\mathcal{E}(w)$ the set of *exact* periods of w, that is,

$$\mathcal{E}(w) = \{d \mid d \in \mathcal{P}(w) \text{ and } d \in \mathcal{D}(n)\}$$

and by $\mathcal{R}(w)$ the reduced set of exact periods of w, that is,

$$\mathcal{R}(w) = \{d \mid d \in \mathcal{E}(w) \text{ and there exists no } d' \in \mathcal{E}(w) \cap \mathcal{D}(d)\}$$

If d is an exact period of w, we set

$$B_d(i) = \{a_j \mid 0 \leq j < n \text{ and } j \equiv i \bmod d\}$$

Now, assume that w is nonprimitive, and let $i_1 < i_2 < \cdots < i_h$ be the elements in $H(w)$. Suppose that w has exact period d and has no free $\diamond$'s with respect to d. Note that for all j, $B_d(i_j) = \{\diamond, b_{i_j}\}$ for some $b_{i_j} \in A$. We define the function f_d as $f_d(i_1, i_2, \ldots, i_h) = (b_{i_1}, b_{i_2}, \ldots, b_{i_h})$. We also define the function f with domain $\mathcal{E}(w)$ where $d \mapsto f_d(i_1, i_2, \ldots, i_h)$, and set $\nu(w) = \|f(\mathcal{E}(w))\|$. Note that $\nu(w)$ is not necessarily equal to $\|\mathcal{E}(w)\|$ since $w = \diamond ba\diamond abab$ is such that $\nu(w) = \|f(\mathcal{E}(w))\| = \|\{(a, b)\}\| = 1$ while $\|\mathcal{E}(w)\| = \|\{2, 4\}\| = 2$.

Lemma 1. *Let w be a nonprimitive partial word that has no free $\diamond$'s with respect to any of its exact periods. Then $\nu(w) = \|\mathcal{R}(w)\|$.*

Proof. Let $i_1 < i_2 < \cdots < i_h$ be the elements in $H(w)$. It is easy to see that $\nu(w) \leq \|\mathcal{R}(w)\|$ because $(i_1, i_2, \ldots, i_h)$ will get mapped to the same h-tuple under both f_d and f_{md} for all integers $m \geq 1$. We now show that f is one-to-one on $\mathcal{R}(w)$. Suppose not, and let $p, q \in \mathcal{R}(w)$ satisfy both $p < q$ and $f_p(i_1, i_2, \ldots, i_h) = f_q(i_1, i_2, \ldots, i_h)$. The bound given by Fine and Wilf's Theorem 1 satisfies $p + q - \gcd(p, q) \leq \frac{n}{3} + \frac{n}{2} - \gcd(p, q) = \frac{5n}{6} - \gcd(p, q) < n$ which implies that any full word of length n with exact periods p, q will also have $\gcd(p, q)$ as an exact period. If we now replace in w the hole in position i_j by b_{i_j} for all j, we obtain a full word w' that has exact periods p, q and thus period $\gcd(p, q)$, and so $\gcd(p, q)$ is also a period of w. Thus $q \in \mathcal{E}(w)$ and $\gcd(p, q) \in \mathcal{E}(w) \cap \mathcal{D}(q)$ implying that $q \notin \mathcal{R}(w)$ which leads to a contradiction. Since f is one-to-one on $\mathcal{R}(w)$, it follows that $\nu(w) \geq \|\mathcal{R}(w)\|$ and thus $\nu(w) = \|\mathcal{R}(w)\|$. □

The following lemma relates to the computation of $\nu(w)$ for nonprimitive pwords w with one hole.

Lemma 2. *If $w \in \mathcal{N}_{1,k}(n)$, then $\nu(w) = 1$.*

Proof. By Lemma 1, since w is a nonprimitive partial word with one hole that has no free $\diamond$'s with respect to any of its exact periods, we have $\nu(w) = \|\mathcal{R}(w)\|$. Now, let $p, q \in \mathcal{R}(w)$ satisfy $p < q$. Since p, q are exact periods, we have $p + q \leq \frac{n}{3} + \frac{n}{2} = \frac{5n}{6} < n$ and Theorem 2 implies that w has $\gcd(p, q)$ as period. But since $q \in \mathcal{E}(w)$, we have that $\gcd(p, q) \in \mathcal{E}(w)$. Since $\gcd(p, q) \neq q$ and $\gcd(p, q) | q$, we get a contradiction with the fact that $q \in \mathcal{R}(w)$. □

We now state some related results.

Lemma 3. *If $w \in \mathcal{N}_{2,k}(n)$, then $\|\mathcal{R}(w)\| = 1$. As a consequence, if n is odd, then $\nu(w) = 1$.*

Proof. Theorem 3 for pwords with two holes gives the optimal bound for the length of w given $p, q \in \mathcal{R}(w)$, that is, $L(2, p, q) = 2p + q - \gcd(p, q)$ with $p < q$. Since p, q are exact periods, we have $p, q \leq \frac{n}{2}$.

- If $p \leq \frac{n}{4}$ and $q \leq \frac{n}{2}$, then $L(2,p,q) \leq \frac{2n}{4} + \frac{n}{2} - \gcd(p,q) = n - \gcd(p,q) < n$.
- If $p = \frac{n}{3}$ and $q = \frac{n}{2}$, then $L(2, \frac{n}{3}, \frac{n}{2}) = \frac{2n}{3} + \frac{n}{2} - \gcd(\frac{2n}{6}, \frac{3n}{6}) = \frac{7n}{6} - \frac{n}{6} = n$.

Thus $L(2,p,q) \leq n$ and Theorem 3 implies that w has $\gcd(p,q)$ as period. Again we get a contradiction as in Lemma 2.

Now, if n is odd, then by Lemma 1, since w is a nonprimitive partial word with two holes that has no free $\diamond$'s with respect to any of its exact periods, we have $\nu(w) = \|\mathcal{R}(w)\|$. □

If $w \in \mathcal{N}_{2,k}(n)$ and n is even and w has two free $\diamond$'s with respect to $d = \frac{n}{2}$, then w has no free $\diamond$'s with respect to any of its exact periods distinct from $\frac{n}{2}$. The smallest exact period being a divisor of $\frac{n}{2}$ by Lemma 3, in such case, we also define $\nu(w) = 1$.

Now, if $w \in \mathcal{N}_{3,k}(n)$, then we can extend the parameter $\nu(w)$ for the case when w has free $\diamond$'s with respect to one of its exact periods d. Again, it is enough to restrain ourselves to the reduced set of exact periods and consider the function f_d for $d \in \mathcal{R}(w)$. The images of $f_d(w)$ and $f_{md}(w)$ are the same for any integer $m > 0$.

In the sequel, given a pword $w \in \mathcal{N}_{h,k}(n)$, the parameter $\nu(w)$ will play an important role. We will obtain words $w' \in \mathcal{N}_{0,k}(n)$ by replacing the $\diamond$'s in w with the corresponding assignments under all possible exact periods of w. Other parameters that will play a role in our counting of primitive partial words are the Φ parameters that we now define. Let v be a pword with h' holes of length d over a k-size alphabet A. Assume that d is a divisor of n and denote by $\mathcal{G}_{n,h,k}(v)$ the set of all nonprimitive pwords w with h holes of length n over A such that there exist pwords $v_0, v_1, \ldots, v_{(\frac{n}{d}-1)}$ and v' satisfying $v_0 = v$, $w = v_0 v_1 \ldots v_{(\frac{n}{d}-1)}$, and $v_i \subset v'$ for all i. If v is nonprimitive (respectively, primitive), then denote by $\phi^N_{n,h,k}(v)$ (respectively, $\phi^P_{n,h,k}(v)$) the number of d''s, nondivisors of d, with $2 \leq d' < d$ such that some $w \in \mathcal{G}_{n,h,k}(v)$ is an $\frac{n}{d'}$-repeat. In other words, $\phi^N_{n,h,k}(v)$ (respectively, $\phi^P_{n,h,k}(v)$) is the number of bases (prefixes) of elements in $\mathcal{G}_{n,h,k}(v)$ shorter than v. Also, let $\Phi^N_{n,h,k}(d,h')$ (respectively, $\Phi^P_{n,h,k}(d,h')$) denote the number of nonprimitive (respectively, primitive) pwords v with h' holes of length d for which $\phi^N_{n,h,k}(v) \geq 1$ (respectively, $\phi^P_{n,h,k}(v) \geq 1$).

We illustrate our ideas with the following example.

Example 1. *First, consider the set*

$$\mathcal{G}_{6,3,2}(a\diamond b) = \{a\diamond b\diamond a\diamond,\ a\diamond b\diamond b\diamond,\ a\diamond b\diamond\diamond b,\ a\diamond ba\diamond\diamond\}$$

Note that any $w \in \mathcal{G}_{6,3,2}(a\diamond b)$ has only 3 as exact period, and thus $\phi^P_{6,3,2}(a\diamond b) = 0$. Now, consider the set $\mathcal{G}_{6,3,2}(ab\diamond) = \{ab\diamond\diamond\diamond a, ab\diamond\diamond\diamond b, ab\diamond\diamond b\diamond, ab\diamond a\diamond\diamond\}$. Note that $ab\diamond\diamond\diamond b$ has exact periods $2, 3$ while the other three words have only 3 as exact period. It follows that $\phi^P_{6,3,2}(ab\diamond) = 1$. The same result holds for pwords in $\{\diamond ab,\ ba\diamond,\ \diamond ba, b\diamond a\}$. Since $\mathcal{P}_{1,2}(3) = \{a\diamond b,\ b\diamond a,\ ab\diamond,\ \diamond ab,\ ba\diamond,\ \diamond ba\}$, we get $\Phi^P_{6,3,2}(3,1) = 4$.

For easy reference, the following table summarizes the notation that will be used in the next sections to count the number of primitive partial words:

Notation	
A	alphabet of size k
k	size of alphabet A
w	partial word of length n over A
n	length of w
d	divisor of n
$\mathcal{D}(n)$	set of divisors of n distinct from n
h	number of holes
$\mathcal{N}_{h,k}(n)$	set of nonprimitive partial words with h holes of length n over A
$N_{h,k}(n)$	cardinality of $\mathcal{N}_{h,k}(n)$
$\mathcal{P}_{h,k}(n)$	set of primitive partial words with h holes of length n over A
$P_{h,k}(n)$	cardinality of $\mathcal{P}_{h,k}(n)$
$\mathcal{T}_{h,k}(n)$	set of partial words with h holes of length n over A
$T_{h,k}(n)$	cardinality of $\mathcal{T}_{h,k}(n)$
$\mathcal{P}(w)$	set of periods of w
$\mathcal{E}(w)$	set of exact periods of w
$\mathcal{R}(w)$	reduced set of exact periods of w

4 Counting primitive partial words: First method

In this section, we first consider all nonprimitive pwords with h holes obtained by replacing h positions in nonprimitive full words with $\diamond$'s, and then subtract the pwords that have been doubly counted. In particular, we express $N_{1,k}(n)$ and $N_{2,k}(n)$ in terms of $N_{0,k}(n)$.

Let $w = a_0a_1 \dots a_{n-1}$ be a full word of length n over an alphabet A of size k. Let $0 \le i_1 < i_2 < \cdots < i_h < n$ and denote by $w_{i_1,\dots,i_h}$ the partial word built from w by replacing positions $i_1, \dots, i_h$ with $\diamond$'s. Setting

$$\mathcal{S}_h(w) = \{w_{i_1,\dots,i_h} \mid 0 \le i_1 < i_2 < \cdots < i_h < n\}$$

we say that w generates each element in the set $\mathcal{S}_h(w)$. For any set X of partial words, we denote by $\mathcal{N}(X)$ the set of nonprimitive pwords in X, that is,

$$\mathcal{N}(X) = \{w \mid w \text{ is nonprimitive and } w \in X\}$$

Lemma 4. *If $w \in \mathcal{N}_{0,k}(n)$, then $\mathcal{S}_h(w) \subset \mathcal{N}_{h,k}(n)$.*

Proof. Since $w \in \mathcal{N}_{0,k}(n)$, there exists a word v such that $w = v^i$ for some $i \ge 2$. If $0 \le i_1 < \cdots < i_h < n$, then $w_{i_1,\dots,i_h} \subset w = v^i$. It follows that $\mathcal{S}_h(w) \subset \mathcal{N}_{h,k}(n)$. □

Denote by $\mathcal{W}_{h,k}(n)$ the set of all nonprimitive partial words with h holes of length n over A obtained by replacing any h positions with $\diamond$'s in nonprimitive full words of length n over A. The following holds:

$$\mathcal{W}_{h,k}(n) = \bigcup_{w \in \mathcal{N}_{0,k}(n)} \mathcal{N}(S_h(w)) = \bigcup_{w \in \mathcal{N}_{0,k}(n)} S_h(w)$$

Obviously,

$$\|\mathcal{W}_{h,k}(n)\| \le \binom{n}{h} N_{0,k}(n)$$

The following lemma states that, given w a full primitive word, the nonprimitive partial word obtained by replacing h positions in w with $\diamond$'s must be in $\mathcal{S}_h(v)$ for some nonprimitive full word v.

Lemma 5. *If $w \in \mathcal{P}_{0,k}(n)$, then $\mathcal{S}_h(w) \subset \mathcal{S}_h(v) \cup \mathcal{P}_{h,k}(n)$ for some $v \in \mathcal{N}_{0,k}(n)$.*

Proof. Let $w \in \mathcal{P}_{0,k}(n)$. If $w_{i_1,\dots,i_h} \in S_h(w)$ is nonprimitive, then there exists a full word u such that $w_{i_1,\dots,i_h} \subset u^i$ for some $i \geq 2$. The word $v = u^i$ is such that $w_{i_1,\dots,i_h} \in S_h(v)$. □

4.1 The one-hole case

The one-hole case is stated in the next theorem.

Theorem 5. *The equality $N_{1,k}(n) = nN_{0,k}(n)$ holds.*

Proof. We first prove by contradiction that if u, v are distinct nonprimitive full words, then $S_1(u) \cap S_1(v) = \emptyset$. Suppose there exists w_i a word of length n such that $w_i \in S_1(u) \cap S_1(v)$. It follows that $w_i = u[0..i)\diamond u[i+1..n)$ and also $w_i = v[0..i)\diamond v[i+1..n)$. Thus $u(j) = v(j)$ for all j such that $0 \leq j < n, j \neq i$. Denote by l_u, l_v the lengths of the primitive roots of u, v and note that $l_u \leq n/2, l_v \leq n/2$. We restrict ourselves to the situation when $i \geq l_u$, otherwise we may just consider the reversals of the words u, v.

Case 1. $l_u = l_v$

It follows that $u(i) = u(i \bmod l_u) = v(i \bmod l_u) = v(i \bmod l_v) = v(i)$, thus $u(i) = v(i)$ which implies that $u = v$, but this contradicts our assumption.

Case 2. $l_u \neq l_v$

Let $d = \gcd(l_u, l_v)$. Since w_i is both l_u- and l_v- periodic, and $l_u + l_v \leq n/2 + n/2 = n$, it follows from Theorem 2 that w_i is also d-periodic. This implies that both u and v must be d-periodic and thus $u(i) = u(i \bmod d) = v(i \bmod d) = v(i)$. It thus follows that $u = v$ which contradicts our assumption.

Now, according to Lemma 4, $\mathcal{W}_{1,k}(n) \subset \mathcal{N}_{1,k}(n)$. By Lemma 5, all nonprimitive pwords obtained by replacing a position with $\diamond$ in a primitive full word can also be derived by replacing a position with $\diamond$ in a nonprimitive full word. In other words, a primitive full word cannot generate a nonprimitive pword which cannot be generated by some nonprimitive full word. By considering both primitive and nonprimitive full words as generators, we have generated all possible partial words with one hole, and of course generated the set $\mathcal{N}_{1,k}(n)$. We can now say that $\mathcal{W}_{1,k}(n) = \mathcal{N}_{1,k}(n)$ and thus

$$N_{1,k}(n) = \|\mathcal{N}_{1,k}(n)\| = \|\mathcal{W}_{1,k}(n)\| = \|\bigcup_{w \in \mathcal{N}_{0,k}(n)} S_1(w)\| = nN_{0,k}(n)$$

□

Corollary 1. *The equality $P_{1,k}(n) = n(P_{0,k}(n) + k^{n-1} - k^n)$ holds.*

Proof. The result follows from the following list of equalities:

$$\begin{aligned} P_{1,k}(n) &= T_{1,k}(n) - N_{1,k}(n) \\ &= T_{1,k}(n) - nN_{0,k}(n) \\ &= T_{1,k}(n) - n(T_{0,k}(n) - P_{0,k}(n)) \\ &= nk^{n-1} - nk^n + nP_{0,k}(n) \\ &= n(P_{0,k}(n) + k^{n-1} - k^n) \end{aligned}$$

□

4.2 The two-hole case

The two-hole case is stated in the next two theorems.

Theorem 6. *For an odd positive integer n, the following equality holds:*

$$N_{2,k}(n) = \binom{n}{2} N_{0,k}(n)$$

Proof. If u, v are distinct nonprimitive full words of length n, then $S_2(u) \cap S_2(v) = \emptyset$. Indeed, suppose there exists a pword $w \in S_2(u) \cap S_2(v)$ such that $u_{i_1,i_2} = w$ and $v_{i_1,i_2} = w$. Thus $u(i) = v(i)$ for all i such that $0 \leq i < n, i \neq i_1, i_2$. Since $w \in \mathcal{N}_{2,k}(n)$ and n is odd, w is not a 2-repeat. It is easy to see that in this case, there are no free $\diamond$'s in w, that is, there exist $j_1, j_2 \in D(w)$ such that $j_1 \equiv i_1 \bmod d$ and $j_2 \equiv i_2 \bmod d$ for each exact period d. This means that in the words u and v, there exists only one pair of assignments for $u(i_1), u(i_2)$ and $v(i_1), v(i_2)$ respectively, since we have already shown in Lemma 3 that $\nu(w) = 1$. It follows that $u(i_1) = v(i_1)$ and $u(i_2) = v(i_2)$ implying that $u = v$ which contradicts our assumption. Similarly to the proof of Theorem 5, since the sets in the union $\bigcup_{w \in \mathcal{N}_{0,k}(n)} S_h(w)$ are pairwise disjoint, we may conclude that

$$N_{2,k}(n) = \|\mathcal{W}_{2,k}(n)\| = \tbinom{n}{2} N_{0,k}(n)$$

□

Theorem 7. *For an even positive integer n, the following equality holds:*

$$N_{2,k}(n) = \binom{n}{2} N_{0,k}(n) - (k-1) T_{1,k}(\frac{n}{2})$$

Proof. It suffices to show that a number of $T_{1,k}(\frac{n}{2})$ words are counted k times each. Let w be a nonprimitive word of even length that generates the partial word $w_{i,j}$. Assume $n \geq 4$. If $w_{i,j}$ is not an $\frac{n}{2}$-repeat, then there are at least three occurrences of the base of length $\leq \frac{n}{3}$. It follows that there are no free $\diamond$'s, which means that the generator w is unique. Assume now that $w_{i,j}$ has an exact period $d = \frac{n}{2}$. If i and j do not belong to the same class modulo d, then again $w_{i,j}$ is uniquely generated since there are no free $\diamond$'s. Now, suppose $i \equiv j \bmod d$ and consider again the pword $w_{i,j}$:

$$\begin{array}{rcccccccc} w_{i,j} & = & a_0 & a_1 & \dots & a_{i-1} & \diamond & a_{i+1} & \dots & a_{d-1} \\ & & a_d & a_{d+1} & \dots & a_{j-1} & \diamond & a_{j+1} & \dots & a_{n-1} \end{array}$$

Note that in this case we have a pair of free $\diamond$'s, which means that in the initial word w, the letter at positions i and j can be any letter in the alphabet, thus a total of k possibilities. The number of partial words u of length $\frac{n}{2}$ with one hole is $T_{1,k}(\frac{n}{2})$. Note that all possible pwords of the form uu can each be generated by k different words in $\mathcal{N}_{0,k}(n)$. Removing $k-1$ copies of such words leaves us with a total of $\frac{1}{2}n(n-1)N_{0,k}(n) - (k-1)T_{1,k}(\frac{n}{2})$, which is what we wanted. □

Corollary 2. *The following holds:*

$$P_{2,k}(n) = \begin{cases} \binom{n}{2}(P_{0,k}(n) + k^{n-2} - k^n) & \text{if } n \text{ is odd} \\ \binom{n}{2}(P_{0,k}(n) + k^{n-2} - k^n) + (k-1)T_{1,k}(\frac{n}{2}) & \text{if } n \text{ is even} \end{cases}$$

Proof. If n is odd, then using Theorem 6 we have the following list of equalities:

$$\begin{aligned} P_{2,k}(n) &= T_{2,k}(n) - N_{2,k}(n) \\ &= T_{2,k}(n) - \binom{n}{2} N_{0,k}(n) \\ &= \binom{n}{2} k^{n-2} - \binom{n}{2} (T_{0,k}(n) - P_{0,k}(n)) \\ &= \binom{n}{2} k^{n-2} - \binom{n}{2} (k^n - P_{0,k}(n)) \\ &= \binom{n}{2} (P_{0,k}(n) + k^{n-2} - k^n) \end{aligned}$$

The case when n is even follows from Theorem 7. □

4.3 The three-hole case

In this section, we discuss the three-hole case. We start with three lemmas (the first holding for any number of holes).

Lemma 6. *If $w \in \mathcal{N}_{h,k}(n)$ and w has no free $\diamond$'s with respect to any of its exact periods, then there exist $\nu(w)$ words in $\mathcal{N}_{0,k}(n)$ that generate w.*

Proof. Let $i_1 < i_2 < \cdots < i_h$ be the elements in $H(w)$. For $p \in \mathcal{R}(w)$, let the h-tuple $(b_{i_1}, \ldots, b_{i_h})$ be the image of $(i_1, \ldots, i_h)$ under f_p (here $b_{i_j} \in A$ for all j). Obviously, replacing for all j the $\diamond$ in position i_j with the corresponding letter b_{i_j} yields a full nonprimitive word that generates w. Since we showed in Lemma 1 that f is bijective on $\mathcal{R}(w)$, it follows that there are $\nu(w) = \|\mathcal{R}(w)\|$ full words that generate w. □

Lemma 7. *If $w \in \mathcal{N}_{3,k}(n)$ and w has three free $\diamond$'s, then there exist $kT_{1,k}(\frac{n}{3})$ words in $\mathcal{N}_{0,k}(n)$ that generate w.*

Proof. The pword w has three free $\diamond$'s only if it is a 3-repeat, that is, $w \subset v^3$ for some pword $v \in \mathcal{T}_{1,k}(\frac{n}{3})$. There also must exist some i, $0 \leq i < \frac{n}{3}$ such that $B_{\frac{n}{3}}(i) = \{\diamond\}$. Let v' denote the full word obtained by replacing the $\diamond$ at position i in v with any letter in A. The resulting full word $(v')^3$ is a generator for w. Since there are k choices to replace the $\diamond$ in v with a letter, it means that k possible full words generate w. Since the total number of words in $\mathcal{N}_{3,k}(n)$ that are 3-repeats is given by $T_{1,k}(\frac{n}{3})$, it follows that $kT_{1,k}(\frac{n}{3})$ words in $\mathcal{N}_{0,k}(n)$ generate w. □

Lemma 8. *If $w \in \mathcal{N}_{3,k}(n)$ and w has two free $\diamond$'s, then there exist $k(n-2)T_{1,k}(\frac{n}{2})$ words in $\mathcal{N}_{0,k}(n)$ that generate w.*

Proof. If w has two free $\diamond$'s, then it must be a 2-repeat, that is, $w \subset v^2$ for some pword $v \in \mathcal{T}_{1,k}(\frac{n}{2})$. There must exist some i, j, $0 \leq i, j < \frac{n}{2}$ such that $B_{\frac{n}{2}}(i) = \{\diamond\}$ and $B_{\frac{n}{2}}(j) = \{\diamond, a\}$ for some $a \in A$. Let v' denote a full word obtained by replacing the $\diamond$'s at positions $i, i + \frac{n}{2}$ within w with any letter in A and the $\diamond$ at position j with the letter a. There are k choices to replace the $\diamond$'s at positions $i, i + \frac{n}{2}$ with a letter. Also, note that there are $(n-2)T_{1,k}(\frac{n}{2})$ words in $\mathcal{N}_{3,k}(n)$ that have a pair of free $\diamond$'s since there are $n-2$ positions where we can place the third $\diamond$. Overall, there are $k(n-2)T_{1,k}(\frac{n}{2})$ words in $\mathcal{N}_{0,k}(n)$ that generate w. □

We have seen that distinct w''s $\in \mathcal{N}_{0,k}(n)$ can generate the same $w \in \mathcal{N}_{3,k}(n)$ whenever w has free $\diamond$'s or $\|\mathcal{R}(w)\| > 1$. Denote by $\mathcal{K}^0_{h,k}(n)$ the set of pwords u with h holes such that u has no free $\diamond$'s and $\|\mathcal{R}(u)\| = 1$. Let $\mathcal{K}^2_{h,k}(n)$ (respectively, $\mathcal{K}^3_{h,k}(n)$) denote the set of pwords u with h holes such that u has two (respectively, three) free $\diamond$'s (such words must have $\|\mathcal{R}(u)\| = 1$). Finally, let $\mathcal{M}_{h,k}(n)$ denote the set of pwords with h holes, no free $\diamond$'s and $\|\mathcal{R}(w)\| = \nu(w) > 1$. We may now conclude the following theorem.

Theorem 8. *The sets $\mathcal{K}^0_{h,k}(n)$, $\mathcal{K}^2_{h,k}(n)$, $\mathcal{K}^3_{h,k}(n)$ and $\mathcal{M}_{h,k}(n)$ represent a partition of $\mathcal{N}_{3,k}(n)$.*

Let us now subtract from $\binom{n}{3}N_{0,k}(n)$ all the words that have been doubly counted. For a set $\mathcal{A}$ we denote by $\pi(\mathcal{A})$ the number of times each word in $\mathcal{A}$ has been counted in $\binom{n}{3}N_{0,k}(n)$.

For n a positive integer, we say that set $\mathcal{X}(n)$ with $\mathcal{X}(n) \subset \mathcal{D}(n)$ is a *θ-set* if $\|\mathcal{X}(n)\| \geq 2$, no element in $\mathcal{X}(n)$ is a multiple of another and $\mathcal{X}(n)$ is maximal. Let $\Psi(n) = \{\mathcal{X}(n) \mid \mathcal{X}(n)$ is a θ-set$\}$.

Example 2. *Computations show that:*

n	$\mathcal{D}(n)$	$\Psi(n)$
6	$\{2,3\}$	$\{\{2,3\}\}$
12	$\{2,3,4,6\}$	$\{\{2,3\},\{3,4\},\{4,6\}\}$
30	$\{2,3,5,6,10,15\}$	$\{\{2,3,5\},\{2,15\},\{3,5\},\{3,10\},\{5,6\},\{6,10,15\}\}$

If $\mathcal{X}(n)$ is a θ-set, then we denote by $\mathcal{M}_{h,k}(n, \mathcal{X}(n))$ the set of pwords w such that $\mathcal{R}(w) = \mathcal{X}(n)$. If $\mathcal{X}_i(n) \neq \mathcal{X}_j(n)$, then $\mathcal{M}_{h,k}(n, \mathcal{X}_i(n)) \cap \mathcal{M}_{h,k}(n, \mathcal{X}_j(n)) = \emptyset$. Finally, let

$$\Omega(n) = \sum_{\mathcal{X}_i(n) \in \Psi(n)} \|\mathcal{M}_{h,k}(n, \mathcal{X}_i(n))\| \|\mathcal{X}_i(n)\|$$

The table below summarizes how many times the elements in each set have been counted in $\binom{n}{3}N_{0,k}(n)$:

$\mathcal{A}$	$\|\mathcal{A}\|$	$\pi(\mathcal{A})$
$\mathcal{K}^0_{h,k}(n)$		1
$\mathcal{K}^2_{h,k}(n)$	$(n-2)T_{1,k}(\frac{n}{2})$	k
$\mathcal{K}^3_{h,k}(n)$	$T_{1,k}(\frac{n}{3})$	k
$\mathcal{M}_{h,k}(n)$		$\Omega(n)$

Theorem 9. *The following equality holds:*

$$N_{3,k}(n) = \binom{n}{3} N_{0,k}(n) - (k-1)(n-2)T_{1,k}(\frac{n}{2}) - (k-1)T_{1,k}(\frac{n}{3}) -$$

$$\sum_{\mathcal{X}_i(n) \in \Psi(n)} \|\mathcal{M}_{h,k}(n, \mathcal{X}_i(n))\| (\|\mathcal{X}_i(n)\| - 1)$$

Whenever $n \not\equiv 0 \bmod j$, we let $T_{1,k}(\frac{n}{j}) = 0$.

5 Counting primitive partial words: Second method

We now count nonprimitive partial words of length n with h holes over a k-size alphabet A through a constructive method. Here, we refine the counting done in the previous sections. Section 5.1 considers the one-hole case, Section 5.2 the two-hole case, and Section 5.3 the three-hole case.

If w is a nonprimitive pword of length n with h holes over A and d is the smallest integer such that there exists a pword v satisfying $w \subset v^{n/d}$, then the *proot* of w is the pword $w[0..d)$. Let $\mathcal{RP}_{h,k}(n, d, h')$ (respectively, $\mathcal{RN}_{h,k}(n, d, h')$) denote the set of nonprimitive pwords of length n with h holes over A with a primitive (respectively, nonprimitive) proot having length d and containing h' holes. Denote by $\mathcal{R}_{h,k}(n, d)$ the set of nonprimitive pwords with h holes of length n over A with a proot of length d. Using the notation $RP_{h,k}(n, d, h') = \|\mathcal{RP}_{h,k}(n, d, h')\|$, $RN_{h,k}(n, d, h') = \|\mathcal{RN}_{h,k}(n, d, h')\|$ and $R_{h,k}(n, d) = \|\mathcal{R}_{h,k}(n, d)\|$, the following equality holds

$$R_{h,k}(n, d) = \sum_{h'=0}^{h} (RP_{h,k}(n, d, h') + RN_{h,k}(n, d, h'))$$

The set $\mathcal{N}_{h,k}(n)$ is generated by considering all possible proots of length $d \in \mathcal{D}(n)$. Different cases occur: The proot belongs to $\mathcal{P}_{h',k}(d)$ for some $h' = 0, \ldots, h$, or the proot belongs to $\mathcal{N}_{h',k}(d)$ for some $h' = 1, \ldots, h$. Note that, in order to avoid double counting, we will never generate nonprimitive pwords starting with a nonprimitive full proot. Therefore, we may always assume that $RN_{h,k}(n, d, 0) = 0$. Given a proot $w[0..d)$ with h' holes, we build the corresponding *temporary* pword $t = (w[0..d))^{n/d}$. We transform t to generate nonprimitive pwords by replacing letters with $\diamond$'s, or vice versa, while the proot remains unchanged. There result pwords containing h holes and having proot $w[0..d)$.

We illustrate the abovementioned ideas with the computation of $N_{1,2}(8)$. Note that the set of lengths of proots is $\{1, 2, 4\}$. The computations below lead to the equality $N_{1,2}(8) = R_{1,2}(8, 1) + R_{1,2}(8, 2) + R_{1,2}(8, 4) = 16 + 16 + 96 = 128$.

If the length of the proot is 1, then $w \subset v^8$. There are $kn = 16$ ways to build such a pword of length 8 with one hole over A, and thus $R_{1,2}(8, 1) = 16$.

Now, if the length of the proot is 2, then $w \subset v^4$. Note that the proot cannot be a primitive partial word with one hole since $P_{1,2}(2) = 0$, and it follows that $RP_{1,2}(8, 2, 1) = 0$. We can therefore split all possible nonprimitive pwords with a proot of length 2 into two sets, based on the nature of the proot. First, if the proot is a primitive full word, then it belongs to the set $\{ab, ba\}$ and we obtain the temporary words $t_1 = abababab$ and $t_2 = babababa$. To obtain nonprimitive pwords with one hole, we replace any of the last six positions in t_1 or t_2 with $\diamond$. Note that replacing any of the first two positions with $\diamond$, thus in the proot, would take us to the previous case, and we would be doubly counting. Six new nonprimitive pwords can be derived from each of the t's and thus $RP_{1,2}(8, 2, 0) = 12$. Second, if the proot is a nonprimitive partial word with one hole, then it belongs to the set $\{a\diamond, b\diamond, \diamond a, \diamond b\}$. There is only one way to build nonprimitive partial words with such proots. For example, if the proot is $a\diamond$, then the only possibility is $a\diamond ababab$. Note that $a\diamond aaaaaa$ is not a possibility since it has already been taken into account. Thus, $RN_{1,2}(8, 2, 1) = 4$ and the equality $R_{1,2}(8, 2) = RP_{1,2}(8, 2, 0) + RP_{1,2}(8, 2, 1) + RN_{1,2}(8, 2, 1) = 16$ holds.

Last, if the length of the proot is 4, then $w \subset v^2$ and again, we split all possible nonprimitive partial words with a proot of length 4 into sets (three sets here), based on the nature of the proot. First, if the proot is a primitive full word, then it belongs to a set of cardinality $P_{0,2}(4) = 12$. To obtain nonprimitive partial words, we may replace any of the last four positions with $\diamond$ and $RP_{1,2}(8, 4, 0) = 48$. Now, if the proot is a primitive partial word with one hole, then it belongs to a set of cardinality $P_{1,2}(4) = 16$. For example, if the proot is $\diamond abb$, then the temporary pword is $t = \diamond abb\diamond abb$. In place of the second $\diamond$, we can put either an a or a b thus obtaining $\diamond abbaabb$ and $\diamond abbbabb$, both nonprimitive partial words. Thus, $RP_{1,2}(8,4,1) = 32$. Finally, if the proot is a nonprimitive partial word with one hole, then it belongs to the set

$$\{aaa\diamond, aa\diamond a, a\diamond aa, \diamond aaa, aba\diamond, ab\diamond b, a\diamond ab, \diamond bab\}$$

unioned with the set containing the pwords obtained by switching a with b, for a total of $N_{1,2}(4) = 16$ proots. There is only one way to build a nonprimitive pword with one hole from such proots. If the proot is $aaa\diamond$, then the only possibility is $aaa\diamond aaab$. Similarly, for the proot $aba\diamond$, the only nonprimitive pword that can be built with this proot is $aba\diamond abaa$. Note that the temporary pword in this case is $t = aba\diamond aba\diamond$, but the second $\diamond$ can be replaced by any letter, except the one letter which will make the pword a 4-repeat with a proot of length 2 (this case has already been taken into account). Thus, $RN_{1,2}(8,4,1) = 16$ and the equality $R_{1,2}(8,4) = RP_{1,2}(8,4,0) + RP_{1,2}(8,4,1) + RN_{1,2}(8,4,1) = 48 + 32 + 16 = 96$ holds.

Note that for two distinct d_1, d_2, we have $\mathcal{R}_{1,2}(8,d_1) \cap \mathcal{R}_{1,2}(8,d_2) = \emptyset$.

The following lemma proves that we are not doubly counting any nonprimitive pwords.

Lemma 9. *Given a proot $w[0..d)$ of length $d \in \mathcal{D}(n)$, the nonprimitive pwords (with one or two holes) generated from $w[0..d)$ have their smallest exact period equal to d.*

Proof. The analysis we are about to perform is similar for the case when we count nonprimitive pwords with two holes. Suppose that during the process of transforming a temporary pword t the resulting pword $w \in \mathcal{N}_{1,k}(n)$ has an exact period d' with $d' < d$. There are four cases we need to consider, depending on whether the proot is a primitive or nonprimitive pword or whether d' is a divisor of d or not. If d' is not a divisor of d, then let $l = \gcd(d, d')$ with $l \neq d'$.

Case 1. $w[0..d)$ is nonprimitive and $d'|d$

This case cannot occur since we subtract the value of $\nu(w[0..d)) = 1$ from the total number of options available to replace the extra $\diamond$'s in t.

Case 2. $w[0..d)$ is nonprimitive and $d' \nmid d$

Since $d, d' \in \mathcal{E}(w)$ it follows from Theorem 2 that $l \in \mathcal{E}(w)$. We are now back in the previous case and no double counting occurs. The reason is that, the pwords w which we are trying to avoid when transforming t have already been avoided in the previous case.

Case 3. $w[0..d)$ is primitive and $d'|d$

This case is easy to deal with simply because of the primitivity of $w[0..d)$. Since w is d'-periodic and $d'|d$, then $w[0..d)$ must also be d'-periodic and thus nonprimitive, which is a contradiction.

Case 4. $w[0..d)$ is primitive and $d' \nmid d$

Since $d, d' \in \mathcal{E}(w)$ it follows from Theorem 2 that $l \in \mathcal{E}(w)$. Since $w = w[0..d)v$ for some pword v and w is d-periodic and l-periodic and $l|d$ it follows that $w[0..d)$ has l as exact period and is thus nonprimitive. This again involves a contradiction with $w[0..d)$ being a primitive pword.

We have now proved that given a proot of length d, the nonprimitive pwords derived from it will always have their smallest exact period equal to d. □

5.1 The one-hole case

The following theorem gives the main result on counting nonprimitive partial words with one hole of length n over A.

Theorem 10. *The following equality holds:*

$$N_{1,k}(n) = kn + \sum_{d|n,\ d\neq 1,n} ((n-d)P_{0,k}(d) + kP_{1,k}(d) + (k-1)N_{1,k}(d)) \quad (1)$$

Proof. Let w be a nonprimitive pword of length n with one hole over A. Let d be the smallest integer such that there exists a pword v satisfying $w \subset v^{n/d}$. Note that $d \in \mathcal{D}(n)$. The case when $d = 1$ can be easily dealt with. There are kn ways we can build a nonprimitive pword of length n with one hole over A, and thus $R_{1,k}(n, 1) = kn$. Consider now the case when $d \in \mathcal{D}(n) \setminus \{1\}$. We split the proof into three cases based on the nature of the proot $w[0..d)$, and we set $t = (w[0..d))^{n/d}$.

Case 1. First, if the proot is a primitive full word, then it belongs to a set of $P_{0,k}(d)$ elements. Transforming t into a nonprimitive pword with one hole requires that we place a $\diamond$ anywhere in t, except in the positions $0, \ldots, d-1$. Since there is a total of $n - d$ such positions, we get

$$RP_{1,k}(n, d, 0) = (n - d)P_{0,k}(d)$$

Case 2. Now, if the proot is a primitive partial word with one hole, then it belongs to a set of $P_{1,k}(d)$ elements. To obtain a nonprimitive pword of length n with one hole, we need to replace in t all the holes, except the first one, with letters in A. Note that once a hole has been replaced with a letter, all remaining holes must be replaced by the same letter. There are k ways we can replace a hole with a letter, thus

$$RP_{1,k}(n, d, 1) = kP_{1,k}(d)$$

Case 3. Finally, if the proot is a nonprimitive partial word with one hole, then it belongs to a set of $N_{1,k}(d)$ elements. Transforming t into a nonprimitive partial word with one hole requires that all holes, except the first one (in the proot), be replaced by a letter in A. Note that once the second hole is replaced by a letter, all remaining holes need to be replaced by the same letter. When replacing the holes, we have all k letters available, except that set of letters that would lead to a nonprimitive pword with one hole and a proot shorter than d, a case that we have already taken into account. Since $\nu(w[0..d)) = 1$, there are $k - \nu(w[0..d)) = k - 1$ nonprimitive partial words with one hole that can be obtained from the temporary pword t above and which have not been counted in previous cases. Since there are $N_{1,k}(d)$ such temporary pwords, it follows that

$$RN_{1,k}(n, d, 1) = (k - 1)N_{1,k}(d)$$

Therefore, the total number of nonprimitive partial words with one hole of length n over A with a proot of length d is

$$\begin{aligned} R_{1,k}(n, d) &= RP_{1,k}(n, d, 0) + RP_{1,k}(n, d, 1) + RN_{1,k}(n, d, 1) \\ &= (n - d)P_{0,k}(d) + kP_{1,k}(d) + (k - 1)N_{1,k}(d) \end{aligned}$$

Denoting by N the right hand side of (1), we want to prove that $N_{1,k}(n) = N$. Note that for a given d, the three cases above cover all possible proots of length d. We do not consider the case of nonprimitive full roots because this falls into the case of full primitive proots with length d' satisfying $d' < d$. Also, once a proot is fixed, we always consider all possible ways the temporary pword t can be transformed into a nonprimitive partial word with one hole, provided we keep the proot unchanged. Modifying the proot by substituting a letter for a $\diamond$ or vice versa, would lead to a nonprimitive word with a different proot (shorter or longer), something that has already been accounted in a different case. We are thus covering all possible nonprimitive partial words with one hole, which implies that $N \geq N_{1,k}(n)$.

We must now prove that $N \leq N_{1,k}(n)$. For a given proot of length d, it holds that the sets $\mathcal{RP}_{1,k}(n, d, 0), \mathcal{RP}_{1,k}(n, d, 1)$ and $\mathcal{RN}_{1,k}(n, d, 1)$ are pairwise disjoint. The reason is that the generating proots for each of the sets are different, as they belong to three different pairwise disjoint sets: $\mathcal{P}_{0,k}(d), \mathcal{P}_{1,k}(d)$ and $\mathcal{N}_{1,k}(d)$. In each of the three cases, the proot is different to start with, and recall that proots remain unchanged throughout the process of transforming a temporary pword into a nonprimitive partial word. Thus, for all three cases, the resulting nonprimitive partial words will be different. Let r_1, r_2 be proots of length d_1, $d_2 \in \mathcal{D}(n)$, with $1 < d_1 < d_2 < n$, and u, v be any pwords such that $u \in \mathcal{R}_{1,k}(n, d_1)$ and $v \in \mathcal{R}_{1,k}(n, d_2)$. Using Lemma 9, u and v have their smallest exact period equal to d_1, respectively d_2. From $d_1 \neq d_2$ it follows that $u \neq v$. Since u, v were any words in $\mathcal{R}_{1,k}(n, d_1)$, respectively $\mathcal{R}_{1,k}(n, d_2)$, it follows that $\mathcal{R}_{1,k}(n, d_1) \cap \mathcal{R}_{1,k}(n, d_2) = \emptyset$. Since d_1, d_2 are any proper divisors of n, it holds that

$$\bigcap_{d_i \in \mathcal{D}(n)} \mathcal{R}_{1,k}(n, d_i) = \emptyset$$

This proves that no double counting occurs and thus $N \leq N_{1,k}(n)$. $\square$

Example 3. *Theorem 10 implies that* $N_{1,2}(8) = 128$ *matching the computations of the previous section. Indeed,*

$$\begin{aligned} N_{1,2}(8) &= 16 + \sum_{d \in \{2,4\}} ((8-d)P_{0,2}(d) + 2P_{1,2}(d) + (2-1)N_{1,2}(d)) \\ &= 16 + 12 + 0 + 4 + 48 + 32 + 16 \\ &= 128 \end{aligned}$$

The formula of Theorem 10 can be further reduced.

Corollary 3. *The equality* $N_{1,k}(n) = nN_{0,k}(n)$ *holds.*

Proof. We prove the equality $N_{1,k}(n) = nN_{0,k}(n)$ by induction on n using Theorem 10. For $n = 1$, the result trivially holds since $N_{1,k}(1) = N_{0,k}(1) = 0$. Assuming the equality holds for all positive integers smaller than n, we get the following sequence of equalities:

$$
\begin{aligned}
N_{1,k}(n) &= kn + \sum_{d|n,\, d\neq 1,n} ((n-d)P_{0,k}(d) + kP_{1,k}(d) + (k-1)N_{1,k}(d)) \\
&= kn + k\sum_{d|n,\, d\neq 1,n} (P_{1,k}(d) + N_{1,k}(d)) + \sum_{d|n, d\neq 1,n} ((n-d)P_{0,k}(d) - N_{1,k}(d)) \\
&= kn + k\sum_{d|n,\, d\neq 1,n} T_{1,k}(d) + \sum_{d|n, d\neq 1,n} ((n-d)P_{0,k}(d) - dN_{0,k}(d)) \\
&= kn + k\sum_{d|n,\, d\neq 1,n} T_{1,k}(d) + n\sum_{d|n, d\neq 1,n} P_{0,k}(d) - \sum_{d|n, d\neq 1,n} (dP_{0,k}(d) + dN_{0,k}(d)) \\
&= kn + k\sum_{d|n,\, d\neq 1,n} dk^{d-1} + n\sum_{d|n, d\neq 1,n} P_{0,k}(d) - \sum_{d|n, d\neq 1,n} dT_{0,k}(d) \\
&= kn + \sum_{d|n,\, d\neq 1,n} dk^{d} + n\sum_{d|n, d\neq 1,n} P_{0,k}(d) - \sum_{d|n, d\neq 1,n} dk^{d} \\
&= kn + n\sum_{d|n,\, d\neq 1,n} P_{0,k}(d) \\
&= n(\sum_{d|n,\, d\neq 1,n} P_{0,k}(d) + k) \\
&= n(\sum_{d|n,\, d\neq 1,n} P_{0,k}(d) + P_{0,k}(1) + P_{0,k}(n) - P_{0,k}(n)) \\
&= n(\sum_{d|n} P_{0,k}(d) - P_{0,k}(n)) \\
&= n(k^n - P_{0,k}(n)) \\
&= n(T_{0,k}(n) - P_{0,k}(n)) \\
&= nN_{0,k}(n)
\end{aligned}
$$

□

5.2 The two-hole case

The following theorem holds.

Theorem 11. *The number of nonprimitive partial words with two holes of length n over a k-size alphabet, $N_{2,k}(n)$, is equal to*

$\binom{n}{2}k + \sum_{d|n,d\neq 1,n}(RP_{2,k}(n,d,0) + RP_{2,k}(n,d,1) + RP_{2,k}(n,d,2) + RN_{2,k}(n,d,1) + RN_{2,k}(n,d,2))$

where

$$RP_{2,k}(n,d,0) = \binom{n-d}{2} P_{0,k}(d) \tag{2}$$

$$RP_{2,k}(n,d,1) = \begin{cases} k(n-d)P_{1,k}(d) & \text{if } d \neq \frac{n}{2} \\ (k(n-d)-(k-1))P_{1,k}(d) & \text{if } d = \frac{n}{2} \end{cases} \tag{3}$$

$$RN_{2,k}(n,d,1) = \begin{cases} (k-1)(n-d)N_{1,k}(d) & \textit{if } d \neq \frac{n}{2} \\ (k-1)(d-1)N_{1,k}(d) & \textit{if } d = \frac{n}{2} \end{cases} \tag{4}$$

$$RP_{2,k}(n,d,2) = k^2 P_{2,k}(d) \tag{5}$$

$$RN_{2,k}(n,d,2) = \begin{cases} (k^2-1)N_{2,k}(d) - (k-1)T_{1,k}(\frac{d}{2}) & \textit{if } d \textit{ is even} \\ (k^2-1)N_{2,k}(d) & \textit{if } d \textit{ is odd} \end{cases} \tag{6}$$

Proof. We give a constructive algorithm for nonprimitive pwords and prove that, along the process of building them, no pwords are missed or double counted.

Let w be a nonprimitive pword of length n with two holes over A, and let d be the smallest integer such that there exists a pword v satisfying $w \subset v^{n/d}$. The case when $d = 1$ can be easily dealt with since there are $\binom{n}{2}k$ ways of building such a nonprimitive pword. Consider now the case when $d \in \mathcal{D}(n) \setminus \{1\}$. We split the proof into five cases based on the nature of the proot $w[0..d) = a_0a_1 \dots a_{d-1}$, and we set $t = (w[0..d))^{n/d}$. If $w[0..d)$ has h' holes, then let $0 \leq i_1 < i_2 < \cdots < i_{h'} < d$ be such that $a_{i_j} = \diamond$. Define

$$C_d = \{l \mid d \leq l < n \text{ and } l \not\equiv i_1 \bmod d, \dots, l \not\equiv i_{h'} \bmod d\}$$
$$D_d(i_j) = \{l \mid d \leq l < n \text{ and } l \equiv i_j \bmod d\} \text{ for all } 1 \leq j \leq h'$$

Note that $\|C_d\| = (\frac{n}{d} - 1)(d - h')$ and $\|D_d(i_j)\| = \frac{n}{d} - 1$. Recall that Lemma 9 guarantees that no double counting will occur.

Case 1. $w[0..d) \in \mathcal{P}_{0,k}(d)$

We need to replace two positions by $\diamond$'s anywhere in t, except in the proot. There is a total of $n - d$ such positions and thus Equality (2) holds.

Case 2. $w[0..d) \in \mathcal{P}_{1,k}(d)$

At this point, all symbols at positions from set $D_d(i_1)$ are $\diamond$'s and all those in C_d are letters. After transforming t into a pword in $\mathcal{N}_{2,h}(n)$, there must remain only one $\diamond$ in the last $n - d$ positions. This can be achieved in two ways, by placing a $\diamond$ in position j with either $j \in C_d$ or $j \in D_d(i_1)$.

Let us first consider the case where $j \in C_d$. There are $\|C_d\|$ options where to place the second $\diamond$ and k choices to pick a letter to replace the positions in $D_d(i_1)$. Note that once a position from set $D_d(i_1)$ has been replaced, all others must be replaced by the same letter. This case yields a total of $k(\frac{n}{d} - 1)(d - 1)$ choices.

Let us now consider the case where $j \in D_d(i_1)$. Note that this can be done in $\|D_d(i_1)\|$ ways and that all positions in C_d remain unchanged. We are now left with $\|D_d(i_1)\| - 1$ $\diamond$'s to be replaced with the same letter. This can be done in k ways provided that $\|D_d(i_1)\| - 1 > 0$, thus a total of $k\|D_d(i_1)\|$ options. If $\|D_d(i_1)\| - 1 = 0$, which implies that $d = n/2$, then this case yields only $\|D_d(i_1)\| = 1$ option, that is $w = w[0..d)w[0..d)$.

Thus, for $d \neq \frac{n}{2}$ we have

$$RP_{2,k}(n,d,1) = (k(\tfrac{n}{d} - 1)(d-1) + k(\tfrac{n}{d} - 1))P_{1,k}(d) = k(n-d)P_{1,k}(d)$$

and if $d = \frac{n}{2}$ then

$$RP_{2,k}(n,d,1) = (k(d-1) + 1)P_{1,k}(d) = (k(n-d) - (k-1))P_{1,k}(d)$$

Putting the two cases together we have that Equality (3) holds.

Case 3. $w[0..d) \in \mathcal{N}_{1,k}(d)$

The approach for this case is similar to the one for Case 2. We replace position j in t with $\diamond$.

Let us first consider the case where $j \in C_d$. There are $\|C_d\|$ options where to place the second $\diamond$, but this time only $k - \nu(w[0..d))$ letters available to replace the $\diamond$'s at positions from $D_d(i_1)$. This last restraint guarantees that the generated pword w will not have an exact period less than d, in other words the proot of w remains unchanged.

Let us now consider the case where $j \in D_d(i_1)$. If $d \neq \frac{n}{2}$, then there are $\|D_d(i_1)\|$ options to place the second $\diamond$ and $k - \nu(w[0..d))$ letters available to replace the remaining $\diamond$'s from positions within set $D_d(i_1)$, thus a total of $(k-\nu(w[0..d)))(\frac{n}{d}-1)$ options. If $d = \frac{n}{2}$, then there is no solution since $w = w[0..d)w[0..d)$ would have a shorter proot.

Thus for $d \neq \frac{n}{2}$, $RN_{2,k}(n,d,1) = ((k-\nu(w[0..d)))(\frac{n}{d}-1)(d-1)+(k-\nu(w[0..d)))(\frac{n}{d}-1))N_{1,k}(d) = (k-\nu(w[0..d)))(n-d)N_{1,k}(d)$. If $d = \frac{n}{2}$, then $RN_{2,k}(n,d,1) = (k-\nu(w[0..d)))(d-1)N_{1,k}(d)$. Keeping in mind that $\nu(w[0..d)) = 1$, we have that Equality (4) holds.

Case 4. $w[0..d) \in \mathcal{P}_{2,k}(d)$

We must replace all positions from the set $D_d(i_1)$ with the same letter, and similarly for $D_d(i_2)$. There are k^2 options to choose these two letters and thus Equality (5) holds.

Case 5. $w[0..d) \in \mathcal{N}_{2,k}(d)$

First, let us consider the case where $w[0..d)$ has a pair of free $\diamond$'s. Of course, this can happen only when $w[0..d)$ is $\frac{d}{2}$-periodic. It is easy to see that the number of pwords $w[0..d)$ of the form $w[0..d) = uu$, where u is a pword with one hole, is equal to $T_{1,k}(d/2)$. We now have k options to replace the $\diamond$'s from positions in set $D_d(i_1)$ and only $k-1$ for the $\diamond$'s from positions within $D_d(i_2)$. The reason why these two letters cannot be the same is because the resulting pword would have a shorter proot, that is u.

Let us now consider the case where $w[0..d)$ does not have a pair of free $\diamond$'s. In this case, we need again to take into account the parameter $\nu(w[0..d))$. We now need to replace all the $\diamond$'s from positions in $D_d(i_1)$ and $D_d(i_2)$ with letters. In order to prevent w from having a proot shorter than d, we must allow only $k^2 - \nu(w[0..d))$ options for choosing the two letters. Since $\nu(w[0..d)) = 1$, we may now conclude that Equality (6) holds.

Note that if we disregard the particular case $d = \frac{n}{2}$, the number of nonprimitive pwords with two holes generated by primitive proots can be summarized as follows:

$$RP_{2,k}(n,d) = \sum_{h'=0}^{2} k^{h'} \binom{n-d}{2-h'} P_{h',k}(d)$$

□

The formula of Theorem 11 can be further reduced.

Corollary 4. *For an odd positive integer n, the following equality holds:*

$$N_{2,k}(n) = \binom{n}{2} N_{0,k}(n)$$

Proof. Setting $n = 2m+1$, we prove the desired equality by induction on m. For $m = 1$, the result trivially holds since $N_{2,k}(3) = \binom{3}{2}k = \binom{3}{2}N_{0,k}(3)$. Assume the equality holds for all positive integers smaller than m. Note that since n is odd, each divisor d of n is odd and so $d \neq \frac{n}{2}$. We have $N_{2,k}(n) =$

$\binom{n}{2}k + \sum_{d|n,\, d\neq 1,n}(\binom{n-d}{2}P_{0,k}(d) + k(n-d)P_{1,k}(d) + (k-1)(n-d)N_{1,k}(d) + k^2 P_{2,k}(d) + (k^2-1)N_{2,k}(d))$

Note that

$$\begin{aligned} k(n-d)P_{1,k}(d) + (k-1)(n-d)N_{1,k}(d) &= k(n-d)T_{1,k}(d) - (n-d)N_{1,k}(d) \\ &= k(n-d)\binom{d}{1}k^{d-1} - (n-d)dN_{0,k}(d) \\ &= (n-d)dk^d - (n-d)dN_{0,k}(d) \\ &= (n-d)dT_{0,k}(d) - (n-d)dN_{0,k}(d) \\ &= (n-d)dP_{0,k}(d) \end{aligned}$$

Similarly

$$\begin{aligned} k^2P_{2,k}(d) + (k^2-1)N_{2,k}(d) &= k^2T_{2,k}(d) - N_{2,k}(d) \\ &= k^2\binom{d}{2}k^{d-2} - \binom{d}{2}N_{0,k}(d) \\ &= \binom{d}{2}k^d - \binom{d}{2}N_{0,k}(d) \\ &= \binom{d}{2}T_{0,k}(d) - \binom{d}{2}N_{0,k}(d) \\ &= \binom{d}{2}P_{0,k}(d) \end{aligned}$$

We hence have the following sequence of equalities:

$$\begin{aligned} N_{2,k}(n) &= \binom{n}{2}k + \sum_{d|n,\ d\neq 1,n}\left(\binom{n-d}{2}P_{0,k}(d) + (n-d)dP_{0,k}(d) + \binom{d}{2}P_{0,k}(d)\right) \\ &= \binom{n}{2}k + \sum_{d|n,\ d\neq 1,n}\binom{n}{2}P_{0,k}(d) \\ &= \binom{n}{2}k + \binom{n}{2}\sum_{d|n,\ d\neq 1,n}P_{0,k}(d) \\ &= \binom{n}{2}(\sum_{d|n,\ d\neq 1,n}P_{0,k}(d) + k) \\ &= \binom{n}{2}(\sum_{d|n,\ d\neq 1,n}P_{0,k}(d) + P_{0,k}(1) + P_{0,k}(n) - P_{0,k}(n)) \\ &= \binom{n}{2}(\sum_{d|n}P_{0,k}(d) - P_{0,k}(n)) \\ &= \binom{n}{2}(k^n - P_{0,k}(n)) \\ &= \binom{n}{2}(T_{0,k}(n) - P_{0,k}(n)) \\ &= \binom{n}{2}N_{0,k}(n) \end{aligned}$$

□

5.3 The three-hole case

Unlike the one- or two-hole case, it may happen that $\nu(w) \neq 1$ when w has three holes. For instance, if $w = ab\diamond\diamond\diamond b$ then $f_2(2,3,4) = (a,b,a)$ and $f_3(2,3,4) = (b,a,b)$, and thus $\nu(w) = 2$. Note that Lemma 9 does not hold anymore for the three-hole case and so we need to consider further modifications to our algorithm. We want to make sure we will not be doubly counting any nonprimitive pwords during the process of generating from a given proot. The parameters $\Phi^N_{n,h,k}(d,h')$ and $\Phi^P_{n,h,k}(d,h')$ will play an important role (they were defined earlier in Section 3). The Φ parameters allow us to avoid generating the word $w = aabaaa\diamond a\diamond a\diamond a$ from the proot $r = aabaaa$. Note that we avoid generating $w = cd\diamond\diamond\diamond d$ from $w[0..d) = cd\diamond$.

Lemma 10. *Given a proot $w[0..d)$ of length $d \in \mathcal{D}(n)$, the nonprimitive pwords with three holes generated from $w[0..d)$ have their smallest exact period equal to d.*

Proof. Suppose that during the process of transforming a temporary pword t the resulting pword $w \in \mathcal{N}_{3,k}(n)$ has an exact period d' with $d' < d$. There are three cases we need to consider, depending on whether the proot is a primitive or nonprimitive pword and whether d' is a divisor of d or not. If d' is not a divisor of d, then let $l = \gcd(d, d')$ with $l \neq d'$.

Case 1. $w[0..d)$ is nonprimitive and $d'|d$

This case cannot occur since we were careful to subtract the value of $\nu(w[0..d))$ from the total number of options available to replace any extra $\diamond$'s in t.

Case 2. $w[0..d)$ is primitive and $d'|d$

Since $w = w[0..d)v$ for some pword v and w is d-periodic and d'-periodic and $d'|d$ it follows that $w[0..d)$ must also be d'-periodic and thus nonprimitive. This again involves a contradiction with $w[0..d)$ being a primitive pword.

Case 3. $d' \nmid d$

We avoid the double counting that may arise from this case by making use of the parameters $\phi^P_{n,h,k}(w[0..d))$ and $\Phi^P_{n,h,k}(d,h')$ and their equivalents for the case when $w[0..d)$ is nonprimitive. From the set of all pwords w generated from a proot $w[0..d)$, we remove those w's which have their smallest exact period shorter than d, that is, the number of pwords satisfying $\phi^P_{n,h,k}(w[0..d)) \geq 1$ (or $\phi^N_{n,h,k}(w[0..d)) \geq 1$). □

Theorem 12. *The number of nonprimitive partial words with three holes of length n over a k-size alphabet, $N_{3,k}(n)$, is equal to*

$$\binom{n}{3}k + \sum_{d|n, d\neq 1,n}(RP_{3,k}(n,d,0) + RP_{3,k}(n,d,1) + RP_{3,k}(n,d,2) + RP_{3,k}(n,d,3) + RN_{3,k}(n,d,1) + RN_{3,k}(n,d,2) + RN_{3,k}(n,d,3))$$

where

$$RP_{3,k}(n,d,0) = \binom{n-d}{3} P_{0,k}(d) - \Phi^P_{n,3,k}(d,0) \tag{7}$$

$$RP_{3,k}(n,d,1) = -\Phi^P_{n,3,k}(d,1) + P_{1,k}(d) \begin{cases} \left(k\binom{(\frac{n}{d}-1)(d-1)}{2} + (\frac{n}{d}-1)(d-1)\right) & \text{if } d = \frac{n}{2} \\ \left(k\binom{(\frac{n}{d}-1)(d-1)}{2} + k(\frac{n}{d}-1)(d-1) + 1\right) & \text{if } d = \frac{n}{3} \\ k\binom{n-d}{2} & \text{otherwise} \end{cases} \tag{8}$$

$$RN_{3,k}(n,d,1) = -\Phi^N_{n,3,k}(d,1) + N_{1,k}(d)\begin{cases} (k-1)\binom{(\frac{n}{d}-1)(d-1)}{2} & \textit{if } d=\frac{n}{2} \\ (k-1)\left(\binom{(\frac{n}{d}-1)(d-1)}{2} + (\frac{n}{d}-1)^2(d-1)\right) & \textit{if } d=\frac{n}{3} \\ (k-1)\binom{n-d}{2} & \textit{otherwise} \end{cases} \tag{9}$$

$$RP_{3,k}(n,d,2) = -\Phi^P_{n,3,k}(d,2) + P_{2,k}(d)\begin{cases} k^2(\frac{n}{d}-1)(d-2)+2k(\frac{n}{d}-1) & \textit{if } d=\frac{n}{2} \\ k^2\binom{n-d}{1} & \textit{if } d\neq\frac{n}{2} \end{cases} \tag{10}$$

$$RN_{3,k}(n,d,2) = -\Phi^N_{n,3,k}(d,2) + \begin{cases} \tau\, T_{1,k}(\frac{d}{2}) + \rho\,(N_{2,k}(d) - T_{1,k}(\frac{d}{2})) & \textit{if } d \textit{ is even} \\ \rho\, N_{2,k}(d) & \textit{if } d \textit{ is odd} \end{cases} \tag{11}$$

$$RP_{3,k}(n,d,3) = k^3 P_{3,k}(d) - \Phi^P_{n,3,k}(d,3) \tag{12}$$

$$RN_{3,k}(n,d,3) = \begin{cases} \lambda(T_{1,k}(\frac{d}{3}) + T_{1,k}(\frac{d}{2})) + \gamma(N_{3,k}(d) - T_{1,k}(\frac{d}{3}) - T_{1,k}(\frac{d}{2})) & \textit{if } d \equiv 0 \bmod 6 \\ \gamma N_{3,k}(d) & \textit{if } d \equiv 1,\,5 \bmod 6 \\ \lambda T_{1,k}(\frac{d}{2}) + \gamma(N_{3,k}(d) - T_{1,k}(\frac{d}{2})) & \textit{if } d \equiv 2,\,4 \bmod 6 \\ \lambda T_{1,k}(\frac{d}{3}) + \gamma(N_{3,k}(d) - T_{1,k}(\frac{d}{3})) & \textit{if } d \equiv 3 \bmod 6 \end{cases} \tag{13}$$

where

$$\tau = \begin{cases} (\frac{n}{d}-1)(d-2)(k^2-k) & \textit{if } d=\frac{n}{2} \\ (n-d)(k^2-k) & \textit{if } d\neq\frac{n}{2} \end{cases}$$

$$\rho = \begin{cases} (k^2-1)(\frac{n}{d}-1)(d-2) + 2(k-1)(\frac{n}{d}-1) & \textit{if } d=\frac{n}{2} \\ (k^2-1)\binom{n-d}{1} & \textit{if } d\neq\frac{n}{2} \end{cases}$$

$$\lambda = k^3 - k$$

$$\gamma = \textstyle\sum_{w[0..d)\in\mathcal{N}_{3,k}(d)}(k^3 - \nu(w[0..d)))$$

Proof. We enumerate all possible options for a generating proot.

Case 1. $w[0..d) \in \mathcal{P}_{0,k}(d)$

Since $w[0..d)$ is a full word, we need to place three $\diamond$'s in the last $n-d$ positions of the word. We thus have Equality (7) holding.

Case 2. $w[0..d) \in \mathcal{P}_{1,k}(d)$

To transform t into a pword in $\mathcal{N}_{3,k}(n)$, we need to place two more $\diamond$'s in the last $n-d$ positions, say i, j, with $d \le i < j < n$.

If we assume $i, j \in C_d$, then there are $k\binom{\|C_d\|}{2}$ pwords that can be obtained, since there are k options to replace the $\diamond$'s currently in positions from $D_d(i_1)$.

Now, assume $i, j \in D_d(i_1)$ (note that this is impossible if $d = \frac{n}{2}$). If $d = \frac{n}{3}$, then $w = w[0..d)w[0..d)w[0..d)$. In general, this case yields $k\binom{\frac{n}{d}-1}{2}$ choices.

Finally, let $i \in C_d$, $j \in D_d(i_1)$ (or vice versa). There are $\|C_d\|$ (respectively, $\|D_d(i_1)\|$) ways to place a $\diamond$ in C_d (respectively, $D_d(i_1)$). If $d \le \frac{n}{3}$, then there are k options to replace the remaining $\diamond$'s at positions from $D_d(i_1)$, which gives a total of $k\|C_d\|\|D_d(i_1)\|$ options. If $d = \frac{n}{2}$, then this case generates only $\|C_d\|$ pwords. We may now conclude that Equality (8) holds. Note that in the case where $d \neq \frac{n}{2}$ and $d \neq \frac{n}{3}$, we have $k\left(\binom{(\frac{n}{d}-1)(d-1)}{2} + (\frac{n}{d}-1)^2(d-1) + \binom{\frac{n}{d}-1}{2}\right) = k\binom{n-d}{2}$.

Case 3. $w[0..d) \in \mathcal{N}_{1,k}(d)$

This case follows the approach of the previous case. Note that by Lemma 2, $\nu(w[0..d)) = 1$.

If we assume $i, j \in C_d$, then this case yields $(k-1)\binom{\|C_d\|}{2}$ options since, in order to keep the pword nonprimitive, only $k-1$ letters are available to replace in t the $\diamond$'s from positions in $D_d(i_1)$.

Now, assume $i, j \in D_d(i_1)$ (note that this is impossible if $d = \frac{n}{2}$ or $d = \frac{n}{3}$). If $d \leq \frac{n}{4}$, then this case yields $(k-1)\binom{\frac{n}{d}-1}{2}$ options.

Finally, let $i \in C_d$, $j \in D_d(i_1)$. Note that this is impossible if $d = \frac{n}{2}$. For $d = \frac{n}{3}$, this case produces $(k-1)(\frac{n}{d}-1)^2(d-1)$ pwords in $\mathcal{N}_{3,k}(n)$. We may now conclude that Equality (9) holds. Note that when $d \neq \frac{n}{2}$ and $d \neq \frac{n}{3}$, we have $(k-1)\Big(\binom{(\frac{n}{d}-1)(d-1)}{2} + (\frac{n}{d}-1)^2(d-1) + \binom{\frac{n}{d}-1}{2}\Big) = (k-1)\binom{n-d}{2}$ pwords.

Case 4. $w[0..d) \in \mathcal{P}_{2,k}(d)$

To transform t, we need to leave t with only one $\diamond$ in some position i with $d \leq i < n$. If $i \in C_d$, then $k^2(\frac{n}{d}-1)(d-2)$ pwords can be generated since the $\diamond$'s from positions within sets $D_d(i_1), D_d(i_2)$ can be replaced by any letter. If $i \in D_d(i_1)$ (or $i \in D_d(i_2)$), then the number of possible pwords this case generates is $k^2(\frac{n}{d}-1)$, unless $d = \frac{n}{2}$ in which case we would have only $k(\frac{n}{d}-1)$ pwords. Overall, Case 4 gives Equality (10). Note that if $d \neq \frac{n}{2}$, then $k^2\big((\frac{n}{d}-1)(d-2) + 2(\frac{n}{d}-1)\big) = k^2\binom{n-d}{1}$.

Case 5. $w[0..d) \in \mathcal{N}_{2,k}(d)$

To transform t into a nonprimitive pword with three holes, we need to leave t with only one hole at position i, with $d \leq i < n$. In this case we need to take into account whether $w[0..d)$ has any free $\diamond$'s or not.

We first consider the case where $w[0..d)$ has two free $\diamond$'s. This implies that $w[0..d)$ must be a 2-repeat and d an even number. There are $T_{1,k}(\frac{d}{2})$ nonprimitive pwords with two holes that are 2-repeats and have a pair of free $\diamond$'s. Denote by τ the number of pwords generated by such a pword with free $\diamond$'s.

If $i \in C_d$, then this case generates $\|C_d\|k(k-1)$ pwords since once the third $\diamond$ has been fixed, the two letters chosen to replace the current $\diamond$'s from positions in $D_d(i_1), D_d(i_2)$ cannot be the same because the resulting pword would have a shorter proot.

If $i \in D_d(i_1)$ and $d \neq \frac{n}{2}$, then this yields $\|D_d(i_1)\|k(k-1)$ possible pwords. It is easy to see that if $d = \frac{n}{2}$, then $\|D_d(i_1)\| = \|D_d(i_2)\| = 1$ and no matter what letter we place at the position in $D_d(i_2)$ the resulting pword would have a shorter proot. The case $i \in D_d(i_2)$ is identical. Note that if $d \neq \frac{n}{2}$, then we have $(\frac{n}{d}-1)(d-2)(k^2-k) + 2(\frac{n}{d}-1)(k^2-k) = (n-d)(k^2-k) = \tau$. Altogether, this case generates τ.

Now, we consider the case where $w[0..d)$ has no free $\diamond$'s. Then the approach is similar to Case 4, the only difference being that once the third $\diamond$ has been fixed, we must subtract $\nu(w[0..d)) = 1$ from the total number of ways we can choose letters to replace the current $\diamond$'s from positions in sets $D_d(i_1), D_d(i_2)$. Thus a proot with no free $\diamond$'s generates ρ and Equality (11) holds.

Case 6. $w[0..d) \in \mathcal{P}_{3,k}(d)$

Since $w[0..d)$ has already three holes, to transform t we need to replace each of the positions from $D_d(i_1)$ with the same letter (and similary for $D_d(i_2)$ and $D_d(i_3)$), a total of k^3 choices, since any combination is allowed. Equality (12) holds.

Case 7. $w[0..d) \in \mathcal{N}_{3,k}(d)$

We have to take into account again the number of free $\diamond$'s $w[0..d)$ has.

If $w[0..d)$ has all three $\diamond$'s free, then $w[0..d)$ is only a 3-repeat. The temporary pword $t = w[0..d)^3$ can generate $k^3 - k$ pwords in $\mathcal{N}_{3,k}(n)$ since we must eliminate the cases when all three letters replacing $\diamond$'s from positions $D_d(i_1), D_d(i_2)$ and $D_d(i_3)$ are the same. Note that there are $T_{1,k}(\frac{d}{3})$ pwords with three free $\diamond$'s.

If $w[0..d)$ has two $\diamond$'s free, then $w[0..d)$ is a 2-repeat and cannot be an m-repeat with $m \geq 3$. Since $w[0..d)$ allows only a 2-repeat, the third $\diamond$ corresponds to one and only one letter, say a_i when mapped through function f. Any combination of three letters can replace the $\diamond$'s from sets $D_d(i_1)$ except the k combinations $\{(a_i, a, a), (a_i, b, b), (a_i, c, c), \ldots\}$. There are $T_{1,k}(\frac{d}{2})$ pwords with two free $\diamond$'s and each can generate $k^3 - k$ pwords in $\mathcal{N}_{3,k}(n)$.

If $w[0..d)$ has no free $\diamond$'s, then each temporary pword t can generate $k^3 - \nu(w[0..d))$ pwords in $\mathcal{N}_{3,k}(n)$. Here different $w[0..d)$'s may have different values for $\nu(w[0..d))$. Equality (13) holds. □

References

[1] J. Berstel and L. Boasson, Partial words and a theorem of Fine and Wilf, *Theoretical Computer Science* **218** (1999) 135–141.

[2] F. Blanchet-Sadri, *Algorithmic Combinatorics on Partial Words* (Chapman & Hall/CRC Press, Boca Raton, FL, 2008).

[3] F. Blanchet-Sadri, Periodicity on partial words, *Computers and Mathematics with Applications* **47** (2004) 71–82.

[4] F. Blanchet-Sadri, Codes, orderings, and partial words, *Theoretical Computer Science* **329** (2004) 177–202.

[5] F. Blanchet-Sadri, Primitive partial words, *Discrete Applied Mathematics* **148** (2005) 195–213.

[6] F. Blanchet-Sadri and A. R. Anavekar, Testing primitivity on partial words, *Discrete Applied Mathematics* **155** (2007) 279–287 (`http://www.uncg.edu/cmp/research/primitive`).

[7] F. Blanchet-Sadri, D. Bal and G. Sisodia, Graph connectivity, partial words, and a theorem of Fine and Wilf, *Information and Computation* **206** (2008) 676–693 (`http://www.uncg.edu/cmp/research/finewilf3`).

[8] F. Blanchet-Sadri, D. Blair and R. V. Lewis, Equations on partial words, *RAIRO-Theoretical Informatics and Applications* **43** (2009) 23–39 (`http://www.uncg.edu/cmp/research/equations`).

[9] F. Blanchet-Sadri and A. Chriscoe, Local periods and binary partial words: an algorithm, *Theoretical Computer Science* **314** (2004) 189–216 (`http://www.uncg.edu/cmp/research/AlgBin`).

[10] F. Blanchet-Sadri and S. Duncan, Partial words and the critical factorization theorem, *Journal of Combinatorial Theory, Series A* **109** (2005) 221–245 (`http://www.uncg.edu/cmp/research/cft`)

[11] F. Blanchet-Sadri, J. Fowler, J. D. Gafni and K. H. Wilson, Combinatorics on partial word correlations, *Journal of Combinatorial Theory, Series A* **117** (2010) 607–624 (`www.uncg.edu/cmp/research/correlations2`).

[12] F. Blanchet-Sadri and R. A. Hegstrom, Partial words and a theorem of Fine and Wilf revisited, *Theoretical Computer Science* **270** (2002) 401–419.

[13] F. Blanchet-Sadri and D. K. Luhmann, Conjugacy on partial words, *Theoretical Computer Science* **289** (2002) 297–312.

[14] C. Choffrut and J. Karhumäki, Combinatorics of Words, in G. Rozenberg and A. Salomaa (Eds.), *Handbook of Formal Languages, Vol. 1, Ch. 6* (Springer-Verlag, Berlin, 1997) 329–438.

[15] M. Crochemore and W. Rytter, *Text Algorithms* (Oxford University Press, 1994).

[16] P. Dömösi, Some results and problems on primitive words, *11th International Conference on Automata and Formal Languages* (2005).

[17] P. Dömösi, S. Horváth and M. Ito, *Context-Free Languages and Primitive Words* (World Scientific, 2011), to appear.

[18] N. J. Fine and H. S. Wilf, Uniqueness theorems for periodic functions, *Proceedings of the American Mathematical Society* **16** (1965) 109–114.

[19] L. J. Guibas and A. M. Odlyzko, Periods in strings, *Journal of Combinatorial Theory, Series A* **30** (1981) 19–42.

[20] M. Lothaire, *Combinatorics on Words* (Addison-Wesley, Reading, MA, 1983; Cambridge University Press, Cambridge, 1997).

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

[22] M. Lothaire, *Applied Combinatorics on Words* (Cambridge University Press, Cambridge, 2005).

[23] R. C. Lyndon and M. P. Schützenberger, The equation $a^m = b^n c^p$ in a free group, *Michigan Mathematical Journal* **9** (1962) 289–298.

[24] H. Petersen, On the language of primitive words, *Theoretical Computer Science* **161** (1996) 141–156.

[25] A. M. Shur and Y. V. Gamzova, Partial words and the interaction property of periods, *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya* **68** (2004) 191–214.

[26] H. J. Shyr and G. Thierrin, Disjunctive languages and codes, Lecture Notes in Computer Science, Vol. 56 (Springer-Verlag, Berlin, 1977) 171–176.

[27] N. J. A. Sloane, *On-Line Encyclopedia of Integer Sequences*, (`http://www.research.att.com/~njas/sequences`).