Discrete Mathematics and Theoretical Computer Science DMTCS vol. **12:3**, 2010, 41–62

On the Number of Balanced Words of Given Length and Height over a Two-Letter Alphabet

Nicolas Bédaride[1] Éric Domenjoud[2] Damien Jamet[2] Jean-Luc Rémy[2]

[1] *Université Aix-Marseille III, Marseille, France.*
[2] *LORIA - Université Nancy 1 - CNRS, Vandœuvre-les-Nancy, France*

received 27 February 2008, *revised 18 April 2010, 7 May 2010*, *accepted 7 May 2010*.

We exhibit a recurrence on the number of discrete line segments joining two integer points in the plane using an encoding of such segments as balanced words of given length and height over the two-letter alphabet $\{0, 1\}$. We give generating functions and study the asymptotic behaviour. As a particular case, we focus on the symmetrical discrete segments which are encoded by balanced palindromes.

Keywords: Symbolic dynamics, complexity function, discrete geometry

1 Introduction

The aim of this paper is to study some properties of discrete lines by using combinatorics on words. The first investigations on discrete lines are dated back to J. Bernoulli[Ber72], E.B. Christoffel [Chr75], A. Markoff [Mar82] and more recently to G.A. Hedlund and H. Morse [MH40] who introduced the terminology of *Sturmian sequences*, for the ones defined on a two-letter alphabet and coding lines with irrational slope. These works gave the first theoretical framework for discrete lines. A sequence $u \in \{0, 1\}^{\mathbb{N}}$ is Sturmian if and only if it is *balanced* and not-eventually periodic. From the 70's, H. Freeman [Fre74], A. Rosenfeld [Ros74] and S. Hung [Hun85] extended these investigations to lines with rational slope and studied *discrete segments*. In [Rev91], J.-P. Reveillès defined arithmetic discrete lines as sets of integer points between two parallel Euclidean lines. There are two sort of arithmetic discrete lines, the *naive* and the *standard* one.

There exists a direct relation between naive (resp. standard) discrete arithmetic lines and Sturmian sequences. Indeed, given a Sturmian sequence $u \in \{0, 1\}^{\mathbb{N}}$, if one associates the letters 0 and 1 with a shifting along the vector $\mathbf{e_1}$ and $\mathbf{e_2}$ (resp. the vectors $\mathbf{e_1}$ and $\mathbf{e_1} + \mathbf{e_2}$) respectively, then, the vertices of the obtained broken line are the ones of a naive arithmetic discrete line (resp. a standard arithmetic discrete line) with the same slope (see Figure 1).

Let $s : \mathbb{N} \longmapsto \mathbb{N}$ be the map defined by:

$$\begin{array}{rcccl} s & : & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & & L & \mapsto & \#\{w \in \{0,1\}^L,\ w \text{ is balanced}\}, \end{array}$$

1365–8050 © 2010 Discrete Mathematics and Theoretical Computer Science (DMTCS), Nancy, France

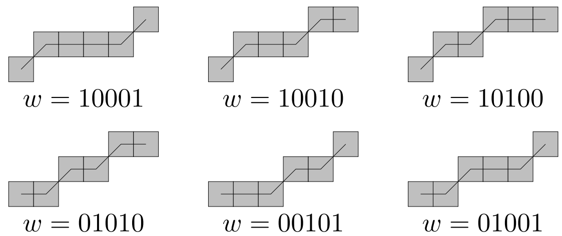

Figure 1: There exist six discrete segments of length 5 and height 2.

where $\#E$ denotes the cardinal of the set E. In other words, given $L \in \mathbb{N}$, $s(L)$ is the number of balanced words of length L, or equivalently, the number of discrete segments of any slope $\alpha \in [0, 1]$ of length L. In [Lip82], it is proved that

$$s(L) = 1 + \sum_{i=1}^{L} (L - i + 1)\varphi(i),$$

where φ is Euler's totient function, that is, $\varphi(n)$ is the number of positive integers smaller that n and coprime with n. Alternative proofs of this result can be found in [Mig91, BP93, CHT02, BL88].

In [dLL05, dLL06], de Luca and De Luca investigated the number $p(L)$ of balanced palindrome words of length $L \in \mathbb{N}$, that is the balanced words coding a symmetrical discrete segments of length L. They proved

$$p(L) = 1 + \sum_{i=0}^{\lceil L/2 \rceil - 1} \varphi(L - 2i).$$

In the present work, we investigate the following question. Given two integer points of $\mathbb{Z}^2$ (also called *pixels* in the discrete geometry literature [CM91]), how many naive discrete segments link these points (see Figure 1)? In other words, given $L \in \mathbb{N}$ and $h \in \mathbb{N}$, how much is $s(L, h) = \#\{w \in \{0, 1\}^L, |w|_1 = h$ and w balanced$\}$? We exhibit a recurrence relation on $s(L, h)$ and generating functions and we study the asymptotic behaviour of the maps s. After this, we focus on the number $p(L, h)$ of balanced palindromes of given length and height for which we also exhibit a recurrence relation and a generating function.

We are interested in these formulas to have a better understanding of the space of Sturmian sequences. Indeed the main combinatorial properties of theses sequences can be seen in similar formulas. For example the formula of $s(L)$ is deeply related to the number of bispecial words of length L, see [CHT02]. One main objective is to generalize these formulas to dimension two in way to understand the combinatorics structure of discrete planes. To a discrete plane is associated a two dimensional word. The study of these words is an interesting problem. The complexity of such a word is not known, the first step in its computation is the following article [DJVV10].

2 Basic notions and notation

Let $\{0, 1\}^*$ and $\{0, 1\}^{\mathbb{N}}$ be the set of respectively finite and infinite words on the alphabet $\{0, 1\}$. We denote the empty word by ϵ. For any word $w \in \{0, 1\}^*$, $|w|$ denotes the length of w, and $|w|_0$ and $|w|_1$ denote respectively the number of 0's and 1's in w. $|w|_1$ is also called the *height* of w. A (finite or infinite) word w is *balanced* if and only if for any finite subwords u and v of w such that $|u| = |v|$, we have

$\big||u|_0 - |v|_0\big| \leq 1$. A (finite or infinite) word w is of type 0 (resp. type 1) if the word w does not contain 11 (resp. the word 00). We denote by $\mathbb{S}$ the set of finite balanced words and by $\mathbb{S}^0$ (resp. $\mathbb{S}^1$) the set of finite balanced words of type 0 (resp. 1).

Let $L, h \in \mathbb{N}$ and $\alpha, \beta \in \{0,1\}^*$. We denote by $\mathbb{S}_{\alpha,\beta}(L,h)$ the set of elements of $\mathbb{S}$ of length L and height h, of which α is a prefix and β is a suffix. Note that α and β may overlap. For short, we usually write $\mathbb{S}(L,h)$ instead of $\mathbb{S}_{\epsilon,\epsilon}(L,h)$. Observe that $\mathbb{S}(L,h)$ is the set of finite balanced words which encode the discrete segments between $(0,0)$ and (L,h). Remark also that $L-h$ is the width of the word, that is the number of zero's. We can count by height or by width, it is the same and this symmetry is used several times in the paper.

We extend the definition of the function $s(L,h)$ on $\mathbb{Z}^2$ by:

$$s(L,h) = \begin{cases} \#\mathbb{S}(L, h \bmod L) & \text{if } L > 0, \\ 1 & \text{if } L = 0 \text{ and } h = 0, \\ 0 & \text{if } L < 0 \text{ or } L = 0 \text{ and } h \neq 0 \end{cases}$$

Observe that for $0 \leq h \leq L$, since $\#\mathbb{S}(L,L) = \#\mathbb{S}(L,0)$, one has $s(L,h) = \#\mathbb{S}(L,h)$.

For $0 \leq h \leq L$ and $\alpha, \beta \in \{0,1\}^*$ we denote by $s_{\alpha,\beta}(L,h)$ the cardinal of $\mathbb{S}_{\alpha,\beta}(L,h)$. Notice that $s_{\alpha,\beta}(L,h) = s_{\overline{\alpha},\overline{\beta}}(L, L-h)$, where $\overline{w}$ is the word obtained by replacing the 0's with 1's and the 1's with 0's in w.

3 General case

3.1 Main theorem

In the present section, we prove the following result:

Theorem 1 *For all $L, h \in \mathbb{N}$ satisfying $0 \leq h \leq L/2$, one has:*

$$s(L,h) = s(L-h-1,h) + s(L-h,h) - s(L-2h-1,h) + s(h-1,L-2) + s(h-1,L-1).$$

In order to prove Theorem 1, let us now introduce some technical definitions and lemmas. Let φ be the morphism defined on $\{0,1\}^*$ and $\{0,1\}^{\mathbb{N}}$ by:

$$\varphi: \begin{array}{ccc} 0 & \mapsto & 0 \\ 1 & \mapsto & 01 \end{array}$$

Let us recall that φ is a Sturmian morphism, that is, for any Sturmian sequence u, the sequence $\varphi(u)$ is Sturmian [Par97, MS93]. Moreover:

Lemma 2 *[Lot02] Let $w \in \{0,1\}^{\mathbb{N}}$.*

1. *If $0w$ is Sturmian of type 0, then there exists a unique Sturmian sequence u satisfying $\varphi(u) = 0w$.*
2. *w is Sturmian if and only if so is $\varphi(w)$.*

Since every balanced word is a factor of a Sturmian word, we directly deduce:

Corollary 3 *If a finite word $w \in \{0,1\}^*$ is balanced then so is $\varphi(w)$.*

Definition 1 (0-erasing map) Let $\theta : \{0,1\}^* \to \{0,1\}^*$ be the map defined by the recurrence relations:

$$\begin{array}{lcll} \theta(\epsilon) & = & \epsilon, & \\ \theta(0^{\alpha+1}) & = & 0^{\alpha} & \text{for } \alpha \geq 0, \\ \theta(1v) & = & 1\theta(v), & \\ \theta(0^{\alpha+1}1v) & = & 0^{\alpha}1\theta(v) & \text{for } \alpha \geq 0, \end{array}$$

Roughly speaking, θ *erases* a 0 in each *maximal range* of 0 in a given word. In some sense, θ is the inverse of φ. Let us now prove some key properties of θ:

Lemma 4 *Consider the set $\mathbb{S}^0_{0,1} = \{u \in \mathbb{S}^0, \exists w \in \{0,1\}^*, u = 0w1\}$ of words in $\mathbb{S}^0$ of the form $0w1$ with $w \in \{0,1\}^*$. Then*

- *The map θ restricted to $\mathbb{S}^0_{0,1}$ is a bijection on $\mathbb{S}_{0,1}$. The map φ restricted to $\mathbb{S}_{0,1}$ is a bijection on $\mathbb{S}^0_{0,1}$.*
- *Moreover we have $\theta(\varphi(w1)) = w1$ for all w.*

Proof: By induction on $|w|_1$.

1. If $|w|_1 = 0$ then $w = 0^{\alpha}$ for some $\alpha \geq 0$ and we have $\varphi(\theta(00^{\alpha}1)) = \varphi(0^{\alpha}1) = 0^{\alpha+1}1 = 0w1$.
2. Assume $|w|_1 \geq 1$ and the result holds for all u such that $|u|_1 < |w|_1$. We have $w = 0^{\alpha}1w'$ for some $\alpha \geq 0$. By assumption, the letter 1 is isolated in $0w1$, so that $w' \neq \epsilon$ and w' starts with the letter 0. Hence,

$$\varphi(\theta(0w1)) = \varphi(\theta(0^{\alpha+1}1w'1)) = \varphi(0^{\alpha}1\theta(w'1)) = 0^{\alpha}01\varphi(\theta(w'1)).$$

By the induction hypothesis, we obtain

$$\varphi(\theta(0w1)) = 0^{\alpha+1}1w'1 = 0w1.$$

□

Example 1 We have by straightforward computations: $\varphi(\theta(11)) = 0101$. Thus the last equation of Lemma 4 is not true everywhere.

Lemma 5 *Let $w \in \{0,1\}^*$. If w is balanced then so is $\theta(w)$.*

Proof:

1. If w is of type 1 (i.e. the letter 0 is isolated in w), then we verify that $\theta(w) = 1^{\alpha}$ for some integer α. Hence it is balanced.
2. Assume now that $w \in \mathbb{S}_0$.

- There exist $\alpha \in \{0,1\}$, $\beta \in \mathbb{N}$ and a Sturmian sequence u of type 0 such that the sequence $0^\alpha w 0^\beta 1 u$ is Sturmian and starts with the letter 0. Notice that u starts with the letter 0 too.
- By point 1 of Lemma 2, there exists a Sturmian sequence u' such that $u = \varphi(u')$.
- We have

$$\begin{array}{rcll} \varphi(\theta(0^\alpha w 0^\beta 1)u') & = & \varphi(\theta(0^\alpha w 0^\beta 1))\varphi(u') & \text{since } \varphi \text{ is a morphism} \\ & = & 0^\alpha w 0^\beta 1 u & \text{by Lemma 4.} \end{array}$$

and by point 2 of Lemma 2, $\theta(0^\alpha w 0^\beta 1)u'$ is Sturmian because $0^\alpha w 0^\beta 1 u$ is Sturmian. Hence $\theta(0^\alpha w 0^\beta 1)$ is balanced as a factor of a balanced word. Finally, we prove by induction on $|w|_1$ that $\theta(0^\alpha w 0^\beta 1) = 0^{\alpha'}\theta(w)0^{\beta'}1$ for some integers α' and β', so that $\theta(w)$ is balanced as a factor of a balanced word.

□

The last technical property of θ we need is:

Lemma 6

1. *If $L \geq 2h+1$ then θ is a bijection from $\mathbb{S}_{0,0}(L,h)$ to $\mathbb{S}_{\epsilon,\epsilon}(L-(h+1),h)$.*
2. *If $L \geq 2h$ then θ is a bijection from $\mathbb{S}_{0,1}(L,h)$ to $\mathbb{S}_{\epsilon,1}(L-h,h)$ and from $\mathbb{S}_{1,0}(L,h)$ to $\mathbb{S}_{1,\epsilon}(L-h,h)$.*
3. *If $L \geq 2h-1$ then θ is a bijection from $\mathbb{S}_{1,1}(L,h)$ to $\mathbb{S}_{1,1}(L-(h-1),h)$.*

Proof: If $h=0$ and $L \neq 0$, $\mathbb{S}_{0,0}(L,h) = \{0^L\}$ and $\mathbb{S}_{\epsilon,\epsilon}(L-(h+1),h) = \{0^{L-1}\} = \{\theta(0^L)\}$. All others sets are empty so that the result obviously holds. In the rest of the proof, we assume $h \geq 1$. We prove the result for $\mathbb{S}_{0,1}(L,h)$. The proof of other cases is similar and left to the reader.

Notice first that $\mathbb{S}_{0,1}(L,h) \subset \mathbb{S}^0$ iff $L \geq 2h$. Indeed, if $L = 2h$, then $\mathbb{S}_{0,1}(L,h) = \{(01)^h\} \subset \mathbb{S}^0$. Now, if $L > 2h$, by the pigeonhole principle, any $w \in \mathbb{S}_{0,1}(L,h)$ must contain the subword 00. Since w is balanced, it cannot contain the subword 11 hence the letter 1 is isolated. Conversely, if the letter 1 is isolated in w, then w must contain at least h 0's, hence $L \geq 2h$.

- $\theta(\mathbb{S}_{0,1}(L,h)) \subset \mathbb{S}_{\epsilon,1}(L-h,h)$

  Let $w \in \mathbb{S}_{0,1}(L,h)$. By an easy induction on h, we show that $|\theta(w)| = L-h$ and $|\theta(w)|_1 = |w|_1 = h$. Furthermore, from Lemma 5, $\theta(w)$ is balanced so that $\theta(w) \in \mathbb{S}(L-h,h)$. Now from the definition of θ, if 1 is a suffix of w, then it is also a suffix of $\theta(w)$ so that $\theta(w) \in \mathbb{S}_{\epsilon,1}(L-h,h)$.

- $\theta : \mathbb{S}_{0,1}(L,h) \to \mathbb{S}_{\epsilon,1}(L-h,h)$ is injective.

  Let $u, v \in \mathbb{S}_{0,1}(L,h)$. We have $u = 0^{\alpha+1}1u'$, $v = 0^{\beta+1}1v'$ and

$$\theta(u) = \theta(v) \Leftrightarrow 0^\alpha 1\theta(u') = 0^\beta 1\theta(v') \Leftrightarrow \alpha = \beta \wedge \theta(u') = \theta(v').$$

  Now, either $h = 1$ and $u' = v' = \epsilon$ so that $u = v$ or $h > 1$ and $u', v' \in \mathbb{S}_{0,1}(L-\alpha-1, h-1)$. We get the result by induction on h.

- $\theta : \mathbb{S}_{0,1}(L,h) \to \mathbb{S}_{\epsilon,1}(L-h,h)$ is surjective.

  Let $w \in \mathbb{S}_{\epsilon,1}(L-h,h)$. We have $w' = \varphi(w) \in \mathbb{S}_{0,1}(L,h)$. Indeed, w' is balanced because w is, $|w'| = |w|_0 + 2|w|_1 = L - 2h + 2h = L$ and $|w'|_1 = |w|_1 = h$ so that $w' \in \mathbb{S}(L,h)$. Since 1 is a suffix of w, it is also a suffix of w' and from Lemma 4, $\theta(w') = w$.

As already said, the proof of other cases is similar. To prove that $\theta : \mathbb{S}_{0,0}(L,h) \to \mathbb{S}_{\epsilon,\epsilon}(L-(h+1),h)$ is surjective, we consider for each $w \in \mathbb{S}_{\epsilon,\epsilon}(L-(h+1),h)$, $w' = \varphi(w)0$. Then we have $w' \in \mathbb{S}_{0,0}(L,h)$ and $\theta(w') = w$. For $\theta : \mathbb{S}_{1,0}(L,h) \to \mathbb{S}_{1,\epsilon}(L-h,h)$, we consider for each $w \in \mathbb{S}_{1,\epsilon}(L-h,h)$, $w' = w''0$ where $\varphi(w) = 0w''$. Finally, for $\theta : \mathbb{S}_{1,1}(L,h) \to \mathbb{S}_{1,1}(L-(h-1),h)$, we consider for each $w \in \mathbb{S}_{1,1}(L-(h-1),h)$, $w' = w''$ where $\varphi(w) = 0w''$.

□

Corollary 7 *For all L,h such that $2 \le h \le L$, $s_{1,1}(L,h) = s_{1,1}(h + (L-h) \bmod (h-1), h)$.*

Proof: Follows from case 3 by induction on $q = \left\lfloor \frac{L-h}{h-1} \right\rfloor$.

□

Lemma 8 *For all L,h such that $0 \le h \le L$, $s_{0,0}(L,h) = s(L-h-1,h)$ and $s_{1,1}(L,h) = s(h-1,L-1)$.*

Proof: We distinguish several cases.

- If $2h < L$, the result is an immediate consequence of case 1 of Lemma 6.
- If $h+1 < L \le 2h$ (which implies $h \ge 2$), we have

$$\begin{array}{rcll}
s_{0,0}(L,h) & = & s_{1,1}(L, L-h) & \text{by exchanging 0's and 1's,} \\
& = & s_{1,1}(L-h+h \bmod (L-h-1), L-h) & \text{by Corollary 7,} \\
& = & s_{0,0}(L-h+h \bmod (L-h-1), h \bmod (L-h-1)) & \text{by exchanging 0's and 1's,} \\
& = & s(L-h-1, h \bmod (L-h-1)) & \text{by case 1 of Lemma 6,} \\
& = & s(L-h-1,h) & \text{by definition of } s(L,h)
\end{array}$$

- If $L = h+1$ we have $L-h-1 = 0$. Either $h = 0$ and $L = 1$ in which case we have $s_{0,0}(1,0) = \#\{0\} = 1 = \#\{\epsilon\} = s(0,0)$, or $h > 0$ and we have $s_{0,0}(L,h) = 0 = s(0,h)$.
- If $L = h$ we have $s_{0,0}(h,h) = 0 = s(-1,h)$.

Since $s_{1,1}(L,h) = s_{0,0}(L,L-h)$, we immediately obtain $s_{1,1}(L,h) = s(h-1,L-h)$. Moreover $s(L,h) = s(L, h \bmod L)$, so that $s_{1,1}(L,h) = s(h-1,L-1)$.

□

We are now ready to prove the main theorem.

Proof of Theorem 1: The property holds for $L = 0$. If $L > 0$, then the following disjoint union holds:

$$\mathbb{S}(L,h) = \mathbb{S}_{0,0}(L,h) \uplus \mathbb{S}_{0,1}(L,h) \uplus \mathbb{S}_{1,0}(L,h) \uplus \mathbb{S}_{1,1}(L,h),$$

and, consequently:

$$s(L,h) = s_{0,0}(L,h) + s_{0,1}(L,h) + s_{1,0}(L,h) + s_{1,1}(L,h).$$

From Lemmas 6 and 8, it follows that

$$\begin{aligned} s(L,h) &= s(L-h-1,h) + s_{\epsilon,1}(L-h,h) + s_{1,\epsilon}(L-h,h) + s(h-1,L-1) \\ &= s(L-h-1,h) + s(L-h,h) + s_{1,1}(L-h,h) - s_{0,0}(L-h,h) + s(h-1,L-1) \\ &= s(L-h-1,h) + s(L-h,h) + s(h-1,L-2) - s(L-2h-1,h) + s(h-1,L-1). \end{aligned}$$

□

3.2 Remark

To summarize, we can compute the formula for $s(L,h)$ for all integers $L,h \in \mathbb{Z}$. Indeed if L is negative, then it is null. If $h = 0$, then $S(L,0) = 1$. If $L = 0$ then $S(0,h) = 1$. The other values can be computed with the statement of Theorem 1 and the relation $s(L,h) = s(L, L-h)$ which is obtained by exchanging 0's and 1's. Sample values of $s(L,h)$ are given in Table 1. The sum of elements in a row give the value of $s(L)$.

$L\backslash h$	0	1	2	3	4	5	6	7	8	9	10
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	4	4	1						
5	1	5	6	6	5	1					
6	1	6	8	6	8	6	1				
7	1	7	11	8	8	11	7	1			
8	1	8	13	12	8	12	13	8	1		
9	1	9	17	13	12	12	13	17	9	1	
10	1	10	20	16	16	10	16	16	20	10	1

Table 1: Sample values of $s(L,h)$ for $0 \le h \le L \le 10$

3.3 An explicit formula for $s(L,2)$

Using the recurrence formula of Theorem 1, we can deduce an explicit formula for some particular cases. Actually we are not able to give an explicit formula in all cases. For instance, one has:

Proposition 9 *Let $L \ge 0$ be an integer. Then, one has:*

$$s(L,2) = \left\lfloor \frac{(L+1)^2+2}{6} \right\rfloor.$$

Proof: By induction on L. One checks that the result holds for $L \in \{0,1,2,3,4\}$. Assume $L \geq 5$ and the result holds for all nonnegative integers smaller than L. From Theorem 1, one deduces $s(L,2) - s(L-3,2) = s(L-2,2) - s(L-5,2) + 2$.

For all $L \geq 3$, let $u_L = s(L,2) - s(L-3,2)$. Then, $u_{L+2} = u_L + 2$ and one obtains $u_L = L - 1 + (L \bmod 2)$. By induction, it follows:

$$s(L,2) = \left\lfloor \frac{(L-2)^2+2}{6} \right\rfloor + L - 1 + (L \bmod 2) = \left\lfloor \frac{L^2+2L}{6} \right\rfloor + (L \bmod 2).$$

Finally, it suffices to check that:

$$\left\lfloor \frac{L^2+2L}{6} + \frac{1}{2} \right\rfloor = \left\lfloor \frac{L^2+2L}{6} \right\rfloor + (L \bmod 2).$$

By considering the remainder of L modulo 6, we obtain that the fractional part of $\frac{L^2+2L}{6}$ is strictly less than $\frac{1}{2}$ if and only if L is even. The result follows. $\square$

3.4 Generating functions of $s(L,h)$

A classical way to obtain an explicit formula of a given function consists in computing its generating function. In this section, we exhibit for each $h \geq 0$ the generating function $\mathcal{S}_h(X)$ of $s(L,h)$, namely $\mathcal{S}_h(X) = \sum_{L \geq 0} s(L,h)X^L$. Let us recall that, in that case, $s(L,h) = \frac{\mathcal{S}_h^{(L)}(0)}{L!}$, where $\mathcal{S}_h^{(L)}$ is the derivative of order L of $\mathcal{S}^{(L)}$.

Theorem 10 *One has:* $\mathcal{S}_0(X) = \frac{1}{1-X}$, $\mathcal{S}_1(X) = \frac{X}{1-X^2}$ *and for all* $h \geq 2$,

$$\mathcal{S}_h(X) = \frac{\mathcal{F}_h(X)}{(1-X^{h-1})(1-X^h)(1-X^{h+1})},$$

where

$$\mathcal{F}_h(X) = (1-X^{h-1})(V_{2h,h}X^h - V_{h-1,h}X^{h+1} - X^{2h-1}) + (1+X)B_h,$$

$$V_{n,h} = \sum_{L=0}^{n-1} s(L,h)X^L,$$

$$B_h = \sum_{r=0}^{h-2} s(h-1,r)X^{r+2h-1}.$$

Proof: We have immediately the two first formulas. From the previous recurrence, for $h \geq 2$, we get:

$$\begin{aligned}
\mathcal{S}_h(X) &= \sum_{L=0}^{2h-1} s(L,h)X^L + \sum_{L\geq 2h} \big(s(L-h-1,h) + s(L-h,h) - s(L-2h-1,h) \\
&\qquad\qquad +s(h-1,L-1) + s(h-1,L-2)\big)X^L
\end{aligned}$$

by using the recurrence of Th. 1.

$$\begin{aligned}
&= \sum_{L=0}^{2h-1} s(L,h)X^L + X^{h+1}\sum_{L\geq h-1} s(L,h)X^L + X^h \sum_{L\geq h} s(L,h)X^L \\
&\quad -X^{2h+1}\sum_{L\geq -1} s(L,h)X^L + \sum_{L\geq 2h}\big(s(h-1,L-1)+s(h-1,L-2)\big)X^L \\
&= \sum_{L=0}^{2h-1} s(L,h)X^L + X^{h+1}\left(\mathcal{S}_h(X) - \sum_{L=0}^{h-2} s(L,h)X^L\right) + X^h\left(\mathcal{S}_h(X) - \sum_{L=0}^{h-1} s(L,h)X^L\right) \\
&\quad -X^{2h+1}\mathcal{S}_h(X) + (X+X^2)\sum_{L\geq 2h-2} s(h-1,L)X^L - s(h-1,0)X^{2h-1} \\
&= (X^h + X^{h+1} - X^{2h+1})\mathcal{S}_h(X) + \sum_{L=0}^{2h-1} s(L,h)X^L - X^{h+1}\sum_{L=0}^{h-2} s(L,h)X^L - X^h\sum_{L=0}^{h-1} s(L,h)X^L \\
&\quad +(X+X^2)\sum_{q\geq 2}\sum_{r=0}^{h-2} s(h-1,r)X^{q\,(h-1)+r} - X^{2h-1}
\end{aligned}$$

by setting $L = q \times (h-1) + r$.

$$\begin{aligned}
&= (X^h + X^{h+1} - X^{2h+1})\mathcal{S}_h(X) + \sum_{L=0}^{2h-1} s(L,h)X^L - X^{h+1}\sum_{L=0}^{h-2} s(L,h)X^L - X^h\sum_{L=0}^{h-1} s(L,h)X^L \\
&\quad +(1+X)\frac{X^{2h-1}}{1-X^{h-1}}\sum_{r=0}^{h-2} s(h-1,r)X^r - X^{2h-1}.
\end{aligned}$$

Finally, we get the formula

$$\mathcal{S}_h(X) \quad = \quad \frac{\mathcal{F}_h(X)}{(1-X^{h-1})(1-X^h)(1-X^{h+1})}$$

where $\mathcal{F}_h \in \mathbb{Z}[X]$ and $\deg(\mathcal{F}_h) \leq 3h-2$. □

Notice that the previous equality does provide a closed formula for $\mathcal{S}_h(X)$ although it still depends on $s(L,h)$ because each sum is finite. Sample values of $\mathcal{S}_h(X)$ are given in Table 2.

3.5 Asymptotic behaviour of $s(L,h)$

Using the generating functions we just computed, we may deduce an expression of $s(L,h)$ which highlights its asymptotic behaviour when L grows.

We prove the following theorem:

$$
\begin{aligned}
\mathcal{S}_2(X) &= \frac{X + X^3}{(1-X)(1-X^2)(1-X^3)} \\
\mathcal{S}_3(X) &= \frac{X + 2X^2 + X^4 + 2X^5}{(1-X^2)(1-X^3)(1-X^4)} \\
\mathcal{S}_4(X) &= \frac{X + X^2 + 3X^3 + 3X^5 + 3X^6 + 3X^7}{(1-X^3)(1-X^4)(1-X^5)} \\
\mathcal{S}_5(X) &= \frac{X + 2X^2 + 3X^3 + 4X^4 + 3X^6 + 5X^7 + 3X^8 + 4X^9 + X^{12}}{(1-X^4)(1-X^5)(1-X^6)} \\
\mathcal{S}_6(X) &= \frac{X + X^2 + X^3 + 4X^4 + 5X^5 + 5X^7 + 10X^8 + 7X^9 + 6X^{10} + 5X^{11} + X^{14}}{(1-X^5)(1-X^6)(1-X^7)}
\end{aligned}
$$

Table 2: Sample values of $\mathcal{S}_h(X)$

Theorem 11 *For all $h \geq 2$, there exist $u_0, \ldots, u_{h-2}, v_0, \ldots, v_{h-1}, w_0, \ldots, w_h \in \mathbb{Q}$ such that*

$$\forall L \geq 0, \quad s(L, h) = \alpha\, L^2 + \beta\, L + u_{L \bmod (h-1)} + v_{L \bmod h} + w_{L \bmod (h+1)}$$

with $\alpha = \dfrac{1}{h\,(h^2-1)} \displaystyle\sum_{i=1}^{h-1} (h-i)\varphi(i)$ and $\beta = \dfrac{1}{h\,(h+1)} \displaystyle\sum_{i=1}^{h} \varphi(i)$ where φ is Euler's totient function.

Before proving this theorem, we need some preliminary results.

Lemma 12 *For all $h \geq 2$, there exist $R, A, B, C \in \mathbb{Q}[X]$ such that $\deg(R) < 3$, $\deg(A) < h-1$, $\deg(B) < h$, $\deg(C) < h+1$ and*

$$\mathcal{S}_h(X) = \frac{R(X)}{(1-X)^3} + \frac{A(X)}{1-X^{h-1}} + \frac{B(X)}{1-X^h} + \frac{C(X)}{1-X^{h+1}}. \tag{1}$$

Proof: We have

$$\mathcal{S}_h(X) = \frac{\mathcal{F}_h(X)}{(1-X^{h-1})(1-X^h)(1-X^{h+1})} = \frac{\mathcal{F}_h(X)}{(1-X)^3\, \dfrac{1-X^{h-1}}{1-X}\, \dfrac{1-X^h}{1-X}\, \dfrac{1-X^{h+1}}{1-X}}$$

where $\mathcal{F}_h \in \mathbb{Z}[X]$ and $\deg(\mathcal{F}_h) \leq 3h-2$.

If h is even, $(1-X)^3$, $\frac{1-X^{h-1}}{1-X}$, $\frac{1-X^h}{1-X}$ and $\frac{1-X^{h+1}}{1-X}$ are pairwise coprime so that

$$\mathcal{S}_h = \frac{R(X)}{(1-X)^3} + \frac{A_0(X)}{\dfrac{1-X^{h-1}}{1-X}} + \frac{B_0(X)}{\dfrac{1-X^h}{1-X}} + \frac{C_0(X)}{\dfrac{1-X^{h+1}}{1-X}}$$

for some $R, A_0, B_0, C_0 \in \mathbb{Q}[X]$ with $\deg(R) < 3$, $\deg(A_0) < h-2$, $\deg(B_0) < h-1$ and $\deg(C_0) < h$. The result follows with $A(X) = (1-X)A_0(X)$, $B(X) = (1-X)B_0(X)$ and $C(X) = (1-X)C_0(X)$.

If h is odd, $\frac{1-X^{h-1}}{1-X}$ and $\frac{1-X^{h+1}}{1-X}$ are divisible by $1+X$. But in this case, we notice that $\mathcal{F}_h(-1) = 0$ so that $\mathcal{F}_h(X)$ is also divisible by $1+X$ and we may write

$$\mathcal{S}_h(X) = \frac{\dfrac{\mathcal{F}_h(X)}{1+X}}{(1-X)^3 \dfrac{1-X^{h-1}}{1-X} \dfrac{1-X^h}{1-X} \dfrac{1-X^{h+1}}{1-X^2}}$$

where $(1-X)^3$, $\frac{1-X^{h-1}}{1-X}$, $\frac{1-X^h}{1-X}$ and $\frac{1-X^{h+1}}{1-X^2}$ are pairwise coprime. Thus we get

$$\mathcal{S}_h = \frac{R(X)}{(1-X)^3} + \frac{A_0(X)}{\dfrac{1-X^{h-1}}{1-X}} + \frac{B_0(X)}{\dfrac{1-X^h}{1-X}} + \frac{C_0(X)}{\dfrac{1-X^{h+1}}{1-X^2}}$$

for some $R, A_0, B_0, C_0 \in \mathbb{Q}[X]$ with $\deg(R) < 3$, $\deg(A_0) < h-2$, $\deg(B_0) < h-1$ and $\deg(C_0) < h-1$. The result follows with $A(X) = (1-X)A_0(X)$, $B(X) = (1-X)B_0(X)$ and $C(X) = (1-X^2)C_0(X)$.

$\square$

Lemma 13 *We recall that $s(L)$ is the number of balanced words of length L, and that it is equal to $\sum_h s(L,h)$. For all $L \geq 0$ we have,*

$$\sum_{h=0}^{L-1} s(L,h)\,h = \frac{L}{2}(s(L)-2).$$

Proof: Let $Z_L = \sum_{h=0}^{L} s(L,h)\,h$. Then

$$\begin{aligned} Z_L &= \sum_{u=0}^{L} s(L,L-u)(L-u) && \text{by setting } h = L-u \\ &= \sum_{u=0}^{L} s(L,u)(L-u) && \text{because } s(L,L-u) = s(L,u) \text{ for all } L,u \in \mathbb{Z} \\ &= L\,s(L) - Z_L \end{aligned}$$

Hence $Z_L = \frac{L}{2}s(L)$ and $\sum_{h=0}^{L-1} s(L,h)\,h = Z_L - L = \frac{L}{2}(s(L)-2)$.

$\square$

Lemma 14 *For all $h \geq 1$,*

$$\sum_{L=h}^{2h-1} s(L,h) - \sum_{L=0}^{h-1} s(L,h) = s(h) + s(h-1) - (h+1)$$

Proof:

$$\begin{aligned}
&\sum_{L=h}^{2h-1} s(L,h) - \sum_{L=0}^{h-1} s(L,h) \\
&= \sum_{L=0}^{h-1} s(L+h,h) - \sum_{L=0}^{h-1} s(L,h) \\
&= \sum_{L=0}^{h-1} (s(h,L) + s(h-1,L) - s(h-L-1,L) + s(L-1,L+h-1) + s(L-1,L+h-2)) - \sum_{L=0}^{h-1} s(L,h) \\
&\qquad \text{by using the recurrence relation from Th. 1} \\
&= \sum_{L=0}^{h-1} (s(h,L) + s(h-1,L) - s(h-1-L,h-1) + s(L-1,h) + s(L-1,h-1)) - \sum_{L=0}^{h-1} s(L,h) \\
&\qquad \text{by using the relation } s(L,h) = s(L, h \bmod L) \text{ on 3rd, 4th and 5th terms} \\
&= s(h) - 1 + s(h-1) - \left(\sum_{u=0}^{h-1} s(u,h-1) - \sum_{L=0}^{h-1} s(L-1,h-1) \right) - \left(\sum_{L=0}^{h-1} s(L,h) - \sum_{L=0}^{h-1} s(L-1,h) \right) \\
&\qquad \text{by setting } u = h-1-L \\
&= s(h) + s(h-1) - (h+1).
\end{aligned}$$

□

We are now ready to prove the main theorem of this section.

Proof of Theorem 11: We first prove existence. In Equation 1, we write

$$R(X) = r_0 + r_1\,(1-X) + r_2\,(1-X)^2.$$

$$A(X) = \sum_{k=0}^{h-2} a_k X^k, \quad B(X) = \sum_{k=0}^{h-1} b_k X^k, \quad C(X) = \sum_{k=0}^{h} c_k X^k.$$

Thus

$$\mathcal{S}_h(X) = \frac{r_0}{(1-X)^3} + \frac{r_1}{(1-X)^2} + \frac{r_2}{1-X} + \frac{A(X)}{1-X^{h-1}} + \frac{B(X)}{1-X^h} + \frac{C(X)}{1-X^{h+1}}.$$

The Taylor expansions of $\frac{1}{(1-X)^3}, \frac{1}{(1-X)^2}, \frac{1}{1-X}$ give the series expansion of $\mathcal{S}_h(X)$:

$$\begin{aligned}
\mathcal{S}_h(X) &= \sum_{L\geq 0} r_0 \, \frac{(L+1)(L+2)}{2} \, X^L + \sum_{L\geq 0} r_1 \, (L+1) \, X^L + \sum_{L\geq 0} r_2 \, X^L \\
&\qquad + \sum_{n\geq 0} A(X) \, X^{n\,(h-1)} + \sum_{n\geq 0} B(X) \, X^{n\,h} + \sum_{n\geq 0} C(X) X^{n\,(h+1)} \\
&= \sum_{L\geq 0} \left(r_0 \, \frac{(L+1)(L+2)}{2} + r_1 \, (L+1) + r_2 \right) X^L \\
&\qquad + \sum_{n\geq 0} \sum_{k=0}^{h-2} a_k X^{n\,(h-1)+k} + \sum_{n\geq 0} \sum_{k=0}^{h-1} b_k X^{n\,h+k} + \sum_{n\geq 0} \sum_{k=0}^{h} c_k X^{n\,(h+1)+k} \\
&= \sum_{L\geq 0} \left(\frac{r_0}{2} \, L^2 + \left(\frac{3}{2} \, r_0 + r_1 \right) \, L + r_0 + r_1 + r_2 + a_{L \bmod (h-1)} + b_{L \bmod h} + c_{L \bmod (h+1)} \right) X^L.
\end{aligned}$$

We get the result with $\alpha = \frac{1}{2}\, r_0$, $\beta = \frac{3}{2}\, r_0 + r_1$, $u_i = a_i + r_0 + r_1 + r_2$, $v_i = b_i$ and $w_i = c_i$.

From the Taylor series of $(1-X)^3 \, \mathcal{S}_h(X)$ at $X = 1$, we get $r_0 = \frac{\mathcal{F}_h(1)}{h\,(h^2-1)}$ and $r_1 = \frac{\frac{3}{2}(h-1)\,\mathcal{F}_h(1) - \mathcal{F}'_h(1)}{h\,(h^2-1)}$.
We have $\mathcal{F}_h(1) = 2 \sum_{r=0}^{h-2} s(h-1,r) = 2(s(h-1)-1)$, and by Lemmas 13 and 14,

$$\begin{aligned}
\mathcal{F}'_h(1) &= -(h-1) \left(\sum_{L=h}^{2\,h-1} s(L,h) - \sum_{L=0}^{h-1} s(L,h) + (h-2) \right) + 2 \sum_{r=0}^{h-2} s(h-1,r) r + (4\,h-1) \sum_{r=0}^{h-2} s(h-1,r) \\
&= -(h-1)(s(h) + s(h-1) - (h+1) + (h-2)) + 2\, \frac{h-1}{2} (s(h-1)-2) + (4h-1)(s(h-1)-1) \\
&= -(h-1)(s(h)-1) + (4h-1)(s(h-1)-1).
\end{aligned}$$

Thus we get finally

$$\begin{aligned}
\alpha &= \frac{\mathcal{F}_h(1)}{2\,h\,(h^2-1)} = \frac{s(h-1)-1}{h\,(h^2-1)} = \frac{1}{h\,(h^2-1)} \sum_{i=1}^{h-1} (h-i)\varphi(i) \\
\beta &= \frac{\frac{3}{2}\, h\, \mathcal{F}_h(1) - \mathcal{F}'_h(1)}{h\,(h^2-1)} = \frac{s(h)-s(h-1)}{h\,(h+1)} = \frac{1}{h\,(h+1)} \sum_{i=1}^{h} \varphi(i).
\end{aligned}$$

□

4 Balanced palindromes and symmetrical discrete segments

In the present section, we focus on the case of segments which are symmetrical with respect to the point $(L/2, h/2)$. These segments are encoded by balanced palindromes.

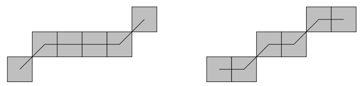

Figure 2: The two symmetrical segments of length 5 and height 2, and their respective encodings as balanced palindromes

4.1 Recurrence formula

This investigation is close to the general case by noticing that if w is a palindrome, then so is $\theta(w)$. We first need an additional property of the mapping θ.

Lemma 15 *For all $w \in \{0,1\}^*$ and all $\alpha \geq 0$,*

$$\begin{array}{rcl}\theta(w1) & = & \theta(w)1 \\ \theta(w10^{\alpha+1}) & = & \theta(w)10^{\alpha}\end{array}$$

Proof: Easy induction on $|w|_1$. □

Corollary 16 *For all $w \in \{0,1\}^*$, $\theta(w\tilde{}) = (\theta(w))\tilde{}$*

Proof: By induction on $|w|_1$. If $|w|_1 = 0$ then $w = 0^\alpha$ for some $\alpha \geq 0$ and the result obviously holds. Now assume that $|w|_1 \geq 1$ and the result holds for all v such that $|v|_1 < |w|_1$. Then either $w = 1v$ and

$$\begin{array}{rcll}\theta(w\tilde{}) & = & \theta((1v)\tilde{}) & \\ & = & \theta(v\tilde{}1) & \\ & = & \theta(v\tilde{})1 & \text{by Lemma 15} \\ & = & \theta(v)\tilde{}1 & \text{by the induction hypothesis} \\ & = & (1\theta(v))\tilde{} & \\ & = & \theta(1v)\tilde{} & \text{by definition of } \theta \\ & = & \theta(w)\tilde{} & \end{array}$$

or $w = 0^{\alpha+1}1v$ and

$$\begin{array}{rcll}\theta(w\tilde{}) & = & \theta((0^{\alpha+1}1v)\tilde{}) & \\ & = & \theta(v\tilde{}10^{\alpha+1}) & \\ & = & \theta(v\tilde{})10^{\alpha} & \text{by Lemma 15} \\ & = & \theta(v)\tilde{}10^{\alpha} & \text{by the induction hypothesis} \\ & = & (0^{\alpha}1\theta(v))\tilde{} & \\ & = & \theta(0^{\alpha+1}1v)\tilde{} & \text{by definition of } \theta \\ & = & \theta(w)\tilde{}. & \end{array}$$

□

Corollary 17 *If $w \in \{0,1\}^*$ is a palindrome, then $\theta(w)$ is also a palindrome.*

Proof: Immediate consequence of Corollary 16 □

We denote by $\mathbb{P}(L,h)$ the set of balanced palindromes of length L and height h, and, for $x \in \{0,1\}$, by $\mathbb{P}_x(L,h)$ the set of balanced palindromes of length L and height h the first (and last) letter of which is x. We define the function $p(L,h)$ on $\mathbb{Z}^2$ by:

$$p(L,h) = \begin{cases} \#\mathbb{P}(L, h \bmod L) & \text{if } L > 0 \\ 1 & \text{if } L = 0 \text{ and } h = 0 \\ 0 & \text{if } L < 0 \text{ or } L = 0 \text{ and } h \neq 0 \end{cases}$$

and for $0 \leq h \leq L$ and $x \in \{0,1\}$, we define $p_x(L,h) = \#\mathbb{P}_x(L,h)$. We have the following properties.

Lemma 18 *Let $L, h \in \mathbb{N}$ such that $0 \leq h \leq L$.*

1. *If $L \geq 2h+1$, then θ is a bijection from $\mathbb{P}_0(L,h)$ to $\mathbb{P}(L-(h+1),h)$.*
2. *If $L \geq 2h-1$, then θ is a bijection from $\mathbb{P}_1(L,h)$ to $\mathbb{P}_1(L-(h-1),h)$.*

Proof:

1. Since $\mathbb{P}_0(L,h) \subset \mathbb{S}_{0,0}(L,h)$, from Lemma 6, we already know that $\theta(\mathbb{P}_0(L,h)) \subset \mathbb{S}(L-(h+1),h)$ and from Corollary 17, $\theta(\mathbb{P}_0(L,h)) \subset \mathbb{P}(L-(h+1),h)$. Since θ is injective on $\mathbb{S}_{0,0}(L,h)$, it is also injective on $\mathbb{P}_0(L,h)$. We are left to prove that it is surjective. Let $w \in \mathbb{P}(L-(h+1),h)$ and $w' = \varphi(w)0$. From Lemma 6, we have $w' \in \mathbb{S}_{0,0}(L,h)$ and $\theta(w') = w$. We prove by induction on $|w|_1$ that w' is a palindrome. If $|w|_1 = 0$ then $w = 0^\alpha$ for some $\alpha \geq 0$ and $w' = 0^{\alpha+1}$ which is trivially a palindrome. If $|w|_1 = 1$ then $w = 0^\alpha 1 0^\alpha$ for some $\alpha \geq 0$ and $w' = 0^{\alpha+1}10^{\alpha+1}$ which is again trivially a palindrome. If $|w|_1 \geq 2$, assume that $\varphi(u)0$ is a palindrome for all u such that $|u|_1 < |w|_1$. We have $w = 0^\alpha 1 v 1 0^\alpha$ for some $\alpha \geq 0$ and some palindrome v with $|v|_1 < |w|_1$. Then $w' = 0^{\alpha+1}1\varphi(v)010^{\alpha+1}$ is a palindrome because $\varphi(v)0$ is a palindrome by the induction hypothesis..
2. The proof is similar. We get in the same way that $\theta(\mathbb{P}_1(L,h)) \subset \mathbb{P}_1(L-(h-1),h)$ and θ is injective on $\mathbb{P}_1(L,h)$. To prove that it is surjective, for each $w \in \mathbb{P}_1(L-(h-1),h)$ we consider w' such that $\varphi(w) = 0w'$. From Lemma 6 we have $\theta(w') = w$ and $w' \in \mathbb{S}_{1,1}(L,h)$. We prove like above that w' is a palindrome so that $w' \in \mathbb{P}_1(L,h)$.

□

Lemma 19 *For all $L, h \in \mathbb{N}$ such that $0 \leq h \leq L$, $p_0(L,h) = p(L-h-1,h)$ and $p_1(L,h) = p(h-1,L-1)$*

Proof: Similar to the proof of Lemma 8. □

From Lemma 19 and the definition of $p(L,h)$, we deduce the following recurrence for $p(L,h)$.

Theorem 20 *Let $L, h \in \mathbb{Z}$,*

$$p(L,h) = \begin{cases} 0 & \textit{if } L < 0 \textit{ or } (L = 0 \textit{ and } h \neq 0) \\ 1 & \textit{if } L \geq 0 \textit{ and } (h = 0 \textit{ or } h = L) \\ p(L, h \bmod L) & \textit{if } L > 0 \textit{ and } (h < 0 \textit{ or } h > L) \\ p(L-h-1,h) + p(h-1,L-1) & \textit{otherwise} \end{cases}$$

Sample values of $p(L,h)$ are given in Table 3.

$L \backslash h$	0	1	2	3	4	5	6	7	8	9	10
0	1										
1	1	1									
2	1	0	1								
3	1	1	1	1							
4	1	0	2	0	1						
5	1	1	2	2	1	1					
6	1	0	2	0	2	0	1				
7	1	1	3	2	2	3	1	1			
8	1	0	3	0	2	0	3	0	1		
9	1	1	3	3	2	2	3	3	1	1	
10	1	0	4	0	2	0	2	0	4	0	1

Table 3: Sample values of $p(L,h)$ for $0 \leq h \leq L \leq 10$

4.2 Generating functions of $p(L,h)$

In the same way we obtained generating functions for $s(L,h)$, we deduce generating functions for $p(L,h)$ from the recurrence above. We consider the generating functions $\mathcal{P}_h(X) = \sum_{L \geq 0} p(L,h) X^L$.

Theorem 21 *One has: $\mathcal{P}_0(X) = \dfrac{1}{1-X}$, $\mathcal{P}_1(X) = \dfrac{X}{1-X^2}$ and for all $h \geq 2$,*

$$\mathcal{P}_h(X) = \frac{1}{1-X^{h+1}} \left(\sum_{L=0}^{h-1} p(L,h) X^L + \frac{X^h}{1-X^{h-1}} \sum_{r=0}^{h-2} p(h-1,r) X^r \right).$$

Proof: One has:

$$\begin{aligned}
\mathcal{P}_0(X) &= \sum_{L\geq 0} p(L,0)X^L = \sum_{L\geq 0} X^L = \frac{1}{1-X}\\
\mathcal{P}_1(X) &= \sum_{L\geq 0} p(L,1)X^L\\
&= p(0,1) + p(1,1)X + \sum_{L\geq 2} p(L,1)X^L\\
&= X + \sum_{L\geq 2}(p(L-2,1) + p(0,L-1))X^L\\
&= X + X^2 \sum_{L\geq 0} p(L,1)X^L\\
&= X + X^2\,\mathcal{P}_1(X).
\end{aligned}$$

Hence, $\mathcal{P}_1(X) = \dfrac{X}{1-X^2}$.

For all $h \geq 2$, we have

$$\begin{aligned}
\mathcal{P}_h(X) &= \sum_{L\geq 0} p(L,h)X^L\\
&= \sum_{L=0}^{h-1} p(L,h)X^L + \sum_{L\geq h} p(L,h)X^L\\
&= \sum_{L=0}^{h-1} p(L,h)X^L + X^h \sum_{L\geq 0} p(L+h,h)X^L\\
&= \sum_{L=0}^{h-1} p(L,h)X^L + X^h \sum_{L\geq 0} (p(L-1,h) + p(h-1,L+h-1))X^L\\
&= \sum_{L=0}^{h-1} p(L,h)X^L + X^h \sum_{L\geq 0} p(L-1,h)X^L + X^h \sum_{L\geq 0} p(h-1,L)X^L\\
&= \sum_{L=0}^{h-1} p(L,h)X^L + X^{h+1}\,\mathcal{P}_h(X) + X^h \sum_{q\geq 0}\sum_{r=0}^{h-2} p(h-1,q(h-1)+r)X^{q(h-1)+r}\\
&= \sum_{L=0}^{h-1} p(L,h)X^L + X^{h+1}\,\mathcal{P}_h(X) + X^h \sum_{q\geq 0} X^{q(h-1)} \sum_{r=0}^{h-2} p(h-1,r)X^r\\
&= X^{h+1}\,\mathcal{P}_h(X) + \sum_{L=0}^{h-1} p(L,h)X^L + \frac{X^h}{1-X^{h-1}} \sum_{r=0}^{h-2} p(h-1,r)X^r.
\end{aligned}$$

Finally, we get

$$\begin{aligned}
\mathcal{P}_h(X) &= \frac{1}{1-X^{h+1}}\left(\sum_{L=0}^{h-1} p(L,h)X^L + \frac{X^h}{1-X^{h-1}}\sum_{r=0}^{h-2} p(h-1,r)X^r\right) \\
&= \frac{(1-X^{h-1})\sum_{L=0}^{h-1} p(L,h)X^L + X^h\sum_{r=0}^{h-2} p(h-1,r)X^r}{(1-X^{h-1})\,(1-X^{h+1})} \\
&= \frac{\mathcal{G}_h(X)}{(1-X^{h-1})\,(1-X^{h+1})}
\end{aligned}$$

where $\mathcal{G}_h(X) \in \mathbb{Z}[X]$ and $\deg(\mathcal{G}_h) \leq 2h-2$. □

Sample values of $\mathcal{P}_h(X)$ are given in Table 4

$$\begin{aligned}
\mathcal{P}_2(X) &= \frac{X}{(1-X)(1-X^3)} & \mathcal{P}_3(X) &= \frac{X}{(1-X^2)(1-X^4)} \\
\mathcal{P}_4(X) &= \frac{X+X^2+X^3}{(1-X^3)(1-X^5)} & \mathcal{P}_5(X) &= \frac{X+X^3+X^7}{(1-X^4)(1-X^6)} \\
\mathcal{P}_6(X) &= \frac{X+X^2+X^3+2\,X^4+X^5+X^8}{(1-X^5)(1-X^7)}
\end{aligned}$$

Table 4: Sample values of $\mathcal{P}_h(X)$

4.3 Asymptotic behaviour of $p(L,h)$

As before, from the generating function, we deduce an expression of $p(L,h)$ which highlights its asymptotic behaviour. We prove the following theorem.

Theorem 22 *For all $h \geq 2$ there exist $u_0,\dots,u_{h-2},v_0,\dots,v_h \in \mathbb{Q}$ such that:*

- *if h is even then*

$$\forall L \geq 0, \quad p(L,h) = \alpha\, L + u_{L \bmod (h-1)} + v_{L \bmod (h+1)}$$

- *if h is odd then*

$$\forall L \geq 0, \quad p(L,h) = \alpha\,(1-(-1)^L)\, L + u_{L \bmod (h-1)} + v_{L \bmod (h+1)}$$

where $\alpha = \dfrac{1}{h^2-1}\displaystyle\sum_{i=1}^{\lceil \frac{h-1}{2} \rceil} \varphi(h+1-2i)$.

Before proving this theorem, we need some lemmas.

Lemma 23 *For all $h \geq 2$:*

- *if h is even then there exist $R, A, B \in \mathbb{Q}[X]$ such that $\deg(R) < 2$, $\deg(A) < h-1$, $\deg(B) < h+1$ and*

$$\mathcal{P}_h(X) = \frac{R(X)}{(1-X)^2} + \frac{A(X)}{1-X^{h-1}} + \frac{B(X)}{1-X^{h+1}}. \tag{2}$$

- *if h is odd then there exist $Q, R, A, B \in \mathbb{Q}[X]$ such that $\deg(Q) < 2$, $\deg(R) < 2$, $\deg(A) < h-1$, $\deg(B) < h+1$ and*

$$\mathcal{P}_h(X) = \frac{Q(X)}{(1+X)^2} + \frac{R(X)}{(1-X)^2} + \frac{A(X)}{1-X^{h-1}} + \frac{B(X)}{1-X^{h+1}} \tag{3}$$

Proof: If h is even then $\mathcal{P}_h(X)$ may be written as

$$\mathcal{P}_h(X) = \frac{\mathcal{G}_h(X)}{(1-X)^2 \, \frac{1-X^{h-1}}{1-X} \, \frac{1-X^{h+1}}{1-X}}.$$

Since $(1-X)^2$, $\frac{1-X^{h-1}}{1-X}$ and $\frac{1-X^{h+1}}{1-X}$ are pairwise coprime, there exist $R, A_0, B_0 \in \mathbb{Q}[X]$ such that $\deg(R) < 2$, $\deg(A_0) < h-2$, $\deg(B_0) < h$ and

$$\mathcal{P}_h(X) = \frac{R(X)}{(1-X)^2} + \frac{A_0(X)}{\frac{1-X^{h-1}}{1-X}} + \frac{B_0(X)}{\frac{1-X^{h+1}}{1-X}}.$$

We get the result with $A(X) = (1-X)A_0(X)$ and $B(X) = (1-X)B_0(X)$.

If h is odd then $1-X^{h-1}$ and $1-X^{h+1}$ are also divisible by $1+X$ so that we may write

$$\mathcal{P}_h(X) = \frac{\mathcal{G}_h(X)}{(1+X)^2 \, (1-X)^2 \, \frac{1-X^{h-1}}{1-X^2} \, \frac{1-X^{h+1}}{1-X^2}}.$$

Since $(1+X)^2$, $(1-X)^2$, $\frac{1-X^{h-1}}{1-X^2}$ and $\frac{1-X^{h+1}}{1-X^2}$ are pairwise coprime, there exist $Q, R, A_0, B_0 \in \mathbb{Q}[X]$ such that $\deg(Q) < 2$, $\deg(R) < 2$, $\deg(A_0) < h-3$, $\deg(B_0) < h-1$ and

$$\mathcal{P}_h(X) = \frac{Q(X)}{(1+X)^2} + \frac{R(X)}{(1-X)^2} + \frac{A_0(X)}{\frac{1-X^{h-1}}{1-X^2}} + \frac{B_0(X)}{\frac{1-X^{h+1}}{1-X^2}}.$$

We get the result with $A(X) = (1-X^2)A_0(X)$ and $B(X) = (1-X^2)B_0(X)$.

□

Lemma 24 *For all $L, h \in \mathbb{N}$ such that L is even and h is odd, $p(L, h) = 0$.*

Proof: By definition of $p(L,h)$, the result is obvious if $L=0$. Also, it is sufficient to prove the result for $0 \leq h \leq L$ because h and $h \bmod L$ have the same parity if L is even. In this case, $p(L,h)$ is exactly the number of balanced palindromes of length L and height h. Let w be a palindrome of length L. If L is even then $w = u\tilde{u}$ for some $u \in \{0,1\}^*$ and $|w|_1 = 2|u|_1$. Hence, there exist no palindrome of even length and odd height. $\square$

We are now ready to prove the main theorem of this section.

Proof of Theorem 22:

The proof is similar to the proof of Theorem 11. In Equations 2 and 3, we write

$$R(X) = \alpha + \beta\,(1-X), \quad Q(X) = \alpha' + \beta'\,(1+X)$$

$$A(X) = \sum_{k=0}^{h-2} a_k X^k, \quad B(X) = \sum_{k=0}^{h} b_k X^k$$

and we get for all $h \geq 2$:

- If h is even then

$$\mathcal{P}_h(X) = \frac{\alpha}{(1-X)^2} + \frac{\beta}{1-X} + \frac{A(X)}{1-X^{h-1}} + \frac{B(X)}{1-X^{h+1}}$$

  and the series expansion of $\mathcal{P}_h(X)$ is

$$\begin{aligned}
\mathcal{P}_h(X) &= \sum_{L\geq 0}(\alpha\,(L+1)+\beta)\,X^L + \sum_{n\geq 0} A(X)\,X^{n(h-1)} + \sum_{n\geq 0} B(X)\,X^{n(h+1)} \\
&= \sum_{L\geq 0}(\alpha\,(L+1)+\beta)\,X^L + \sum_{n\geq 0}\sum_{k=0}^{h-2} a_k X^{n(h-1)+k} + \sum_{n\geq 0}\sum_{k=0}^{h} b_k X^{n(h+1)+k} \\
&= \sum_{L\geq 0}(\alpha\,L + \alpha + \beta + a_{L \bmod (h-1)} + b_{L \bmod (h+1)})\,X^L
\end{aligned}$$

  Hence

$$\forall L \geq 0,\ p(L,h) = \alpha\,L + \alpha + \beta + a_{L \bmod (h-1)} + b_{L \bmod (h+1)}$$

- If h is odd then

$$\mathcal{P}_h(X) = \frac{\alpha'}{(1+X)^2} + \frac{\beta'}{1+X} + \frac{\alpha}{(1-X)^2} + \frac{\beta}{1-X} + \frac{A(X)}{1-X^{h-1}} + \frac{B(X)}{1-X^{h+1}}$$

  and its series expansion is

$$\mathcal{P}_h(X) = \sum_{L\geq 0}(\alpha'\,(-1)^L\,(L+1) + \beta'\,(-1)^L + \alpha\,(L+1) + \beta + a_{L \bmod (h-1)} + b_{L \bmod (h+1)})X^L.$$

  Hence

$$\forall L \geq 0,\ p(L,h) = \alpha'\,(-1)^L\,(L+1) + \beta'\,(-1)^L + \alpha\,(L+1) + \beta + a_{L \bmod (h-1)} + b_{L \bmod (h+1)}.$$

Considering the Taylor expansion of $(1-X)^2\mathcal{P}_h(X)$ at $X=1$ and, in case h is even, $(1+X)^2\mathcal{P}_h(X)$ at $X=-1$, we get

$$\alpha = \frac{\mathcal{G}_h(1)}{h^2-1} = \frac{1}{h^2-1}\sum_{r=0}^{h-2} p(h-1,r) = \frac{1}{h^2-1}\sum_{i=1}^{\lceil \frac{h-1}{2} \rceil} \varphi(h+1-2i)$$

and, if h is odd,

$$\alpha' = \frac{\mathcal{G}_h(-1)}{h^2-1} = \frac{(-1)^h}{h^2-1}\sum_{r=0}^{h-2}(-1)^r p(h-1,r) = \frac{-1}{h^2-1}\sum_{r=0}^{h-2} p(h-1,r)$$

where the last equality is deduced from Lemma 24. Hence $\alpha' = -\alpha$.

Finally, we get the result with $u_i = \alpha+\beta+a_i$ if h is even and $u_i = (1-(-1)^i)\alpha+\beta+(-1)^i\beta'+a_i$ if h is odd and $v_i = b_i$.

□

Acknowledgements

The authors would like to thank Julien Cassaigne for useful discussions and comments.

References

[Ber72] J. Bernoulli. Sur une nouvelle espèce de calcul. In *Recueil pour les Astronomes*, volume T. 1, pages 255–284, Berlin, 1772.

[BL88] C. A. Berenstein and D. Lavine. On the number of digital straight line segments. *P.A.M.I*, 10(6):880–887, 1988.

[BP93] J. Berstel and M. Pocchiola. A geometric proof of the enumeration formula for Sturmian words. *Int. J. Algebra Comput.*, 3(3):349–355, 1993.

[Chr75] E. B. Christoffel. Observatio arithmetica. *Math. Ann.*, 6:145–152, 1875.

[CHT02] J. Cassaigne, P. Hubert, and S. Troubetzkoy. Complexity and growth for polygonal billiards. *Annales de l'Institut Fourier*, 52:1001–1013, 2002. 12 pages, 4 figures.

[CM91] J.-M. Chassery and A. Montanvert. *Géométrie discrète en analyse d'images*. Hermès, Paris, 1991.

[DJVV10] E. Domenjoud, D. Jamet, D. Vergnaud, and L. Vuillon. On the number of local configuration of size 2xn on discrete planes. *Preprint*, 2010.

[dLL05] A. de Luca and A. De Luca. Palindromes in sturmian words. In Clelia de Felice and Antonio Restivo, editors, *Developments in Language Theory*, volume 3572 of *Lecture Notes in Computer Science*, pages 199–208. Springer, 2005.

[dLL06] A. de Luca and A. De Luca. Combinatorial properties of Sturmian palindromes. *Internat. J. Found. Comput. Sci.*, 17(3):557–573, 2006.

[Fre74] H. Freeman. Computer processing of line-drawing images. *ACM Computing Surveys*, 6(1):57–97, 1974.

[Hun85] S. H. Y. Hung. On the straightness of digital arcs. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume PAMI-7, pages 203–215, 1985.

[Lip82] E. P. Lipatov. On a classification of binary collections and properties of homogeneity classes. *Problemy Kibernet.*, 39:67–84, 1982.

[Lot02] M. Lothaire. *Algebraic Combinatorics On Words*. Cambridge University Press, 2002.

[Mar82] A. Markoff. Sur une question de Jean Bernoulli. *Math. Ann.*, 19:27–36, 1882.

[MH40] M. Morse and G. A. Hedlund. Symbolic dynamics II. Sturmian trajectories. *Amer J. Math.*, 62:1–42, 1940.

[Mig91] F. Mignosi. On the number of factors of sturmian words. *Theoret. Comput. Sci.*, 82(1):71–84, 1991.

[MS93] Filippo Mignosi and Patrice Séébold. Morphismes sturmiens et règles de Rauzy. *J. Théor. Nombres Bordeaux*, 5(2):221–233, 1993.

[Par97] B. Parvaix. Propriétés d'invariance des mots sturmiens. *J. Théor. Nombres Bordeaux*, 9:351–369, 1997.

[Rev91] J.-P. Reveillès. *Géométrie discrète, calcul en nombres entiers et algorithmique*. Thèse d'Ètat, Université Louis Pasteur, Strasbourg, 1991.

[Ros74] A. Rosenfeld. Digital straight lines segments. In *IEEE Transactions on Computers*, volume 12, pages 1264–1369, 1974.