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

Nicolas Bédaride

Université Aix-Marseille III
Avenue de l'Escadrille Normandie-Niémen
13397 Marseille Cedex 20, France
email : Nicolas.Bedaride@univ.u-3mrs.fr

Eric Domenjoud, Damien Jamet, Jean-Luc Rémy

ADAGIo Team – LORIA
BP 239
F- 54506 Vandoeuvre-lès-Nancy Cedex
email : {Eric.Domenoud, Damien.Jamet, Jean-Luc.Remy}@loria.fr

Words 2007 – Marseille, sep. 2007

Context : discrete geometry for computer graphics

Definition (Arithmetical discrete Lines (J.-P. Reveilles, 1991))

Let $\mathbf{v} = (v_1, v_2) \in \mathbb{Z}^2$ with $\gcd(v_1, v_2) = 1$, $\mu \in \mathbb{Z}$ and $\theta \in \mathbb{Z}$. The arithmetical discrete line $D(\mathbf{v}, \mu, \theta)$ with normal vector $\mathbf{v}$, shift μ and arithmetical thickness θ is the subset of $\mathbb{Z}^2$ defined by

$$D(\mathbf{v}, \mu, \theta) = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^2 \text{ and } 0 \leq \langle \mathbf{x}, \mathbf{v} \rangle - \mu < \theta\}$$

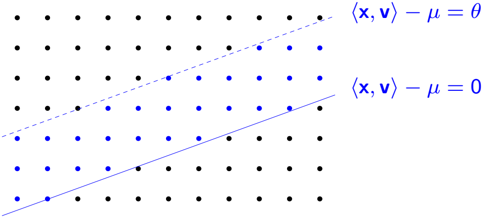


FIG.: general discrete line in $\mathbb{Z}^2$

Context : discrete geometry for computer graphics

Definition (Arithmetical discrete Lines (J.-P. Reveilles, 1991))

Let $\mathbf{v} = (v_1, v_2) \in \mathbb{Z}^2$ with $\gcd(v_1, v_2) = 1$, $\mu \in \mathbb{Z}$ and $\theta \in \mathbb{Z}$. The arithmetical discrete line $D(\mathbf{v}, \mu, \theta)$ with normal vector $\mathbf{v}$, shift μ and arithmetical thickness θ is the subset of $\mathbb{Z}^2$ defined by

$$D(\mathbf{v}, \mu, \theta) = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^2 \text{ and } 0 \leq \langle \mathbf{x}, \mathbf{v} \rangle - \mu < \theta\}$$

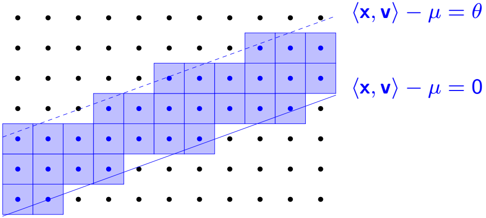


FIG.: general discrete line in $\mathbb{Z}^2$

Discrete lines are usually represented using unit squares called *pixels*.

Context : discrete geometry for computer graphics

Definition (Arithmetical discrete Lines (J.-P. Reveilles, 1991))

Let $\mathbf{v} = (v_1, v_2) \in \mathbb{Z}^2$ with $\gcd(v_1, v_2) = 1$, $\mu \in \mathbb{Z}$ and $\theta \in \mathbb{Z}$. The arithmetical discrete line $D(\mathbf{v}, \mu, \theta)$ with normal vector $\mathbf{v}$, shift μ and arithmetical thickness θ is the subset of $\mathbb{Z}^2$ defined by

$$D(\mathbf{v}, \mu, \theta) = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^2 \text{ and } 0 \leq \langle \mathbf{x}, \mathbf{v} \rangle - \mu < \theta\}$$

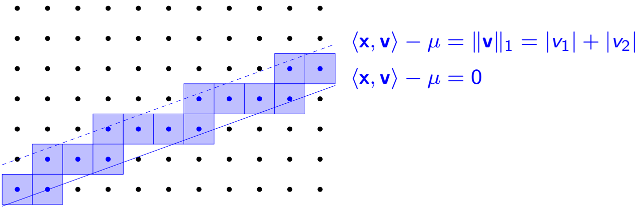


FIG.: *standard* discrete line : $\theta = \|\mathbf{v}\|_1$

The standard discrete line is the thinest 1-connected discrete line.

Context : discrete geometry for computer graphics

Definition (Arithmetical discrete Lines (J.-P. Reveilles, 1991))

Let $\mathbf{v} = (v_1, v_2) \in \mathbb{Z}^2$ with $\gcd(v_1, v_2) = 1$, $\mu \in \mathbb{Z}$ and $\theta \in \mathbb{Z}$. The arithmetical discrete line $D(\mathbf{v}, \mu, \theta)$ with normal vector $\mathbf{v}$, shift μ and arithmetical thickness θ is the subset of $\mathbb{Z}^2$ defined by

$$D(\mathbf{v}, \mu, \theta) = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^2 \text{ and } 0 \leq \langle \mathbf{x}, \mathbf{v} \rangle - \mu < \theta\}$$

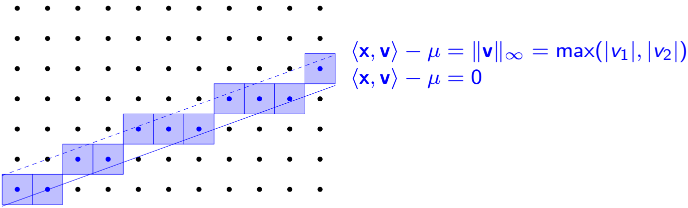


FIG.: *naive* discrete line : $\theta = \|\mathbf{v}\|_\infty$

The naive discrete line is the thinest (0-)connected discrete line.

Context : discrete geometry for computer graphics

Definition (Arithmetical discrete Lines (J.-P. Reveilles, 1991))

Let $\mathbf{v} = (v_1, v_2) \in \mathbb{Z}^2$ with $\gcd(v_1, v_2) = 1$, $\mu \in \mathbb{Z}$ and $\theta \in \mathbb{Z}$. The arithmetical discrete line $D(\mathbf{v}, \mu, \theta)$ with normal vector $\mathbf{v}$, shift μ and arithmetical thickness θ is the subset of $\mathbb{Z}^2$ defined by

$$D(\mathbf{v}, \mu, \theta) = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^2 \text{ and } 0 \leq \langle \mathbf{x}, \mathbf{v} \rangle - \mu < \theta\}$$

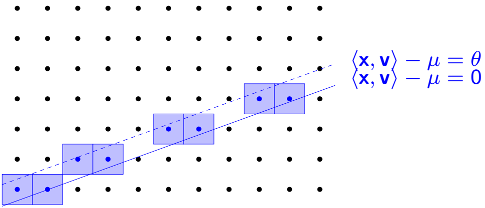


FIG.: disconnected discrete line : $\theta < \|\mathbf{v}\|_\infty$

Context : discrete geometry for computer graphics

Definition (Arithmetical discrete Lines (J.-P. Reveilles, 1991))

Let $\mathbf{v} = (v_1, v_2) \in \mathbb{Z}^2$ with $\gcd(v_1, v_2) = 1$, $\mu \in \mathbb{Z}$ and $\theta \in \mathbb{Z}$. The arithmetical discrete line $D(\mathbf{v}, \mu, \theta)$ with normal vector $\mathbf{v}$, shift μ and arithmetical thickness θ is the subset of $\mathbb{Z}^2$ defined by

$$D(\mathbf{v}, \mu, \theta) = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^2 \text{ and } 0 \leq \langle \mathbf{x}, \mathbf{v} \rangle - \mu < \theta\}$$

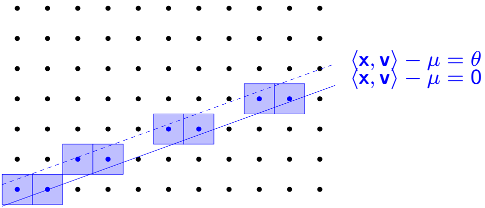


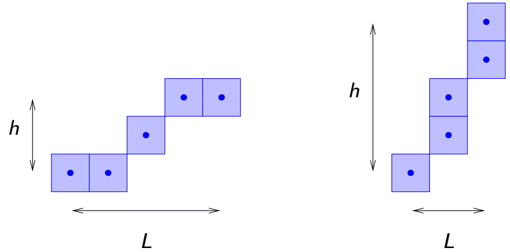
FIG.: disconnected discrete line : $\theta < \|\mathbf{v}\|_\infty$

In this talk, we consider only *naive* discrete lines

Discrete segments

Definition

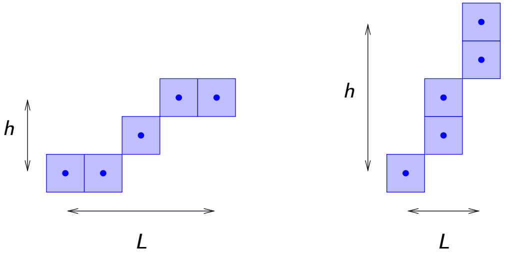
A (naive) discrete segment is a finite connected subset of a (naive) discrete line.



Discrete segments

Definition

A (naive) discrete segment is a finite connected subset of a (naive) discrete line.



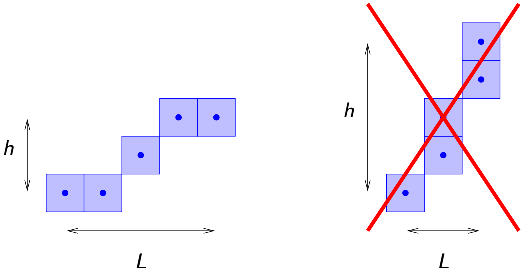
Questions

1. How many discrete segments exist of a given length L?
2. How many discrete segments exist of a given length L and height h?

Discrete segments

Definition

A (naive) discrete segment is a finite connected subset of a (naive) discrete line.



Questions

1. How many discrete segments exist of a given length L?
2. How many discrete segments exist of a given length L and height h?

For symmetry reasons, we consider only segments with $0 \leq h \leq L$

Question

How many discrete segments exist of a given length L?

Theorem (F. Mignosi, 1991)

The number $s(L)$ of discrete segments of length L is

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

where φ is Euler totient function.

Example

$s(L) = 1, 2, 4, 8, 14, 24, 36, 54, 76, 104 \ldots$

Question

How many discrete segments exist of a given length L and height h ?

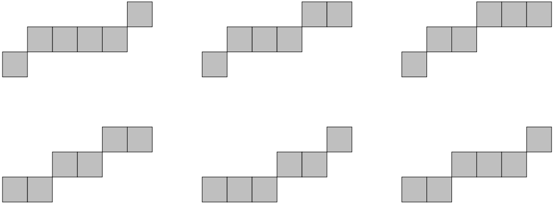


FIG.: The 6 discrete segments of length 5 and height 2

Question

How many discrete segments exist of a given length L and height h?

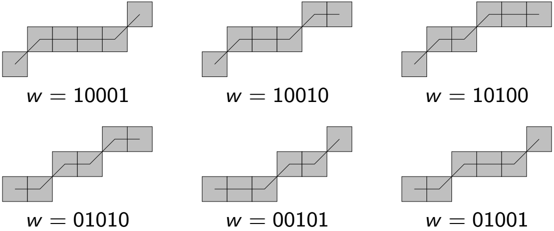


Fig.: The 6 discrete segments of length 5 and height 2 and the corresponding encodings

Discrete segments are encoded by finite *balanced* words.

Definition (Balanced words)

A word w over the alphabet $\{0,1\}$ is balanced iff for all subwords u and v of w,

$$|u| = |v| \implies \big||u|_1 - |v|_1\big| \leq 1.$$

Properties

- *A balanced word never contains both the subwords 00 and 11*
  $\implies$ *in a balanced word, at least one letter (0 or 1) is isolated.*
- *A word w encodes a discrete segment iff w is balanced.*

$$\underbrace{\dots 010 \dots 010 \dots\dots 010 \dots 010 \dots}_{\text{blocks of 0's separated by isolated 1's}}$$

$$\underbrace{\dots 101 \dots 101 \dots\dots 101 \dots 101 \dots}_{\text{blocks of 1's separated by isolated 0's}}$$

Question

How many balanced words exist of a given length L and height h?

Notations

- Let $\mathbb{S}$ denote the set of finite balanced words over the alphabet $\{0,1\}$.
- For all $L, h \in \mathbb{N}$ such that $0 \le h \le L$,

$$S(L,h) = \{w \mid w \in \mathbb{S},\ |w| = L \text{ and } |w|_1 = h\}$$

and for all $L, h \in \mathbb{Z}$,

$$s(L,h) = \begin{cases} |S(L, h \bmod L)| & \text{if } L \ge 1 \\ 1 & \text{if } L = 0 \text{ and } h = 0 \\ 0 & \text{if } L < 0 \text{ or } (L = 0 \text{ and } h \ne 0) \end{cases}$$

- For all $L, h \in \mathbb{N}$ such that $0 \le h \le L$ and all $x, y \in \{0,1\}^*$,

$$S_{x,y}(L,h) = \{w \mid w \in S(L,h) \text{ and } \exists u, v \in \{0,1\}^\star, w = xu = vy\}$$

and

$$s_{x,y}(L,h) = |S_{x,y}(L,h)|.$$

$$s(L,h) = s_{\epsilon,\epsilon}(L,h) = \begin{cases} 1 & \text{if } L = h = 0, \\ s_{0,0}(L,h) + s_{0,1}(L,h) + s_{1,0}(L,h) + s_{1,1}(L,h) & \text{if } L > 0. \end{cases}$$

Lemma (Symmetry)

If $w \in S(L,h)$ *and* $h > L/2$ *then* $\overline{w} \in S(L, L-h)$ *and we have* $L - h < L/2$

$\Longrightarrow$ We have $s(L,h) = s(L, L-h)$ and we may restrict to the case $h \leq L/2$, i.e. to the words w which contain more 0's than 1's (1's are isolated).

Definition (Step contraction)

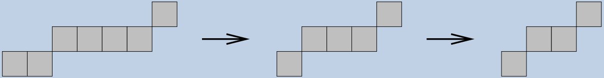
Let $w \in S(L,h)$, $w = 0^{\lambda_0}10^{\lambda_1}1\ldots10^{\lambda_{h-1}}10^{\lambda_h}$ with $\lambda_0, \lambda_h \geq 0$ et $\lambda_1, \ldots, \lambda_{h-1} \geq 1$:

$$\theta_0(w) = 0^{\lambda'_0}10^{\lambda_1-1}1\ldots10^{\lambda_{h-1}-1}10^{\lambda'_h}$$

with $\lambda'_i = \max(0, \lambda_i - 1)$
(θ_0 *erases* a 0 in each maximal block of 0's)

Example

$w = 010001$ $\quad$ $\theta_0(w) = 1001$ $\quad$ $\theta_0^2(w) = 101$



Lemma

If w *is balanced, then so is* $\theta_0(w)$.

Lemma

If $h > L/2$ then

- θ_0 *is a bijection from* $S_{0,0}(L,h)$ *to* $S_{\epsilon,\epsilon}(L-(h+1),h)$
- θ_0 *is a bijection from* $S_{0,1}(L,h)$ *to* $S_{\epsilon,1}(L-h,h)$ *and from* $S_{1,0}(L,h)$ *to* $S_{1,\epsilon}(L-h,h)$
- θ_0 *is a bijection from* $S_{1,1}(L,h)$ *to* $S_{1,1}(L-(h-1),h)$

Idea of the proof

Let $w \in S_{0,1}(L,h)$

$$w = \underbrace{\underbrace{0\ldots0}_{\geq 1}1\underbrace{0\ldots0}_{\geq 1}10\ldots\ldots01\underbrace{0\ldots0}_{\geq 1}1}_{h \text{ blocks of 0's}}$$

- θ_0 *erases* h 0's $\Rightarrow |\theta_0(w)| = L - h \Rightarrow \theta_0(w) \in S(L-h,h)$
- The number of 0's at the beginning of w is ≥ 1
  $\Rightarrow$ The number of 0's at the beginning of $\theta_0(w)$ is ≥ 0
  $\Rightarrow$ The first letter of $\theta_0(w)$ is unknown but the last one is 1
  $\Rightarrow \theta_0(w) \in S_{\epsilon,1}(L-h,h)$
- Injection is obvious, surjection a bit less.

Other cases are similar.

Lemma

If $h > L/2$ then

- θ_0 *is a bijection from* $S_{0,0}(L, h)$ *to* $S_{\epsilon,\epsilon}(L - (h + 1), h)$
- θ_0 *is a bijection from* $S_{0,1}(L, h)$ *to* $S_{\epsilon,1}(L - h, h)$ *and from* $S_{1,0}(L, h)$ *to* $S_{1,\epsilon}(L - h, h)$
- θ_0 *is a bijection from* $S_{1,1}(L, h)$ *to* $S_{1,1}(L - (h - 1), h)$

Corollary

If $L > 2h$ then

- $s_{0,0}(L, h) = s_{\epsilon,\epsilon}(L - (h + 1), h)$
- $s_{0,1}(L, h) = s_{\epsilon,1}(L - h, h)$
- $s_{1,0}(L, h) = s_{1,\epsilon}(L - h, h)$
- $s_{1,1}(L, h) = s_{1,1}(L - (h - 1), h)$

Recurrence formula

Theorem

For all $L, h \in \mathbb{Z}^2$,

$$s(L,h) = \begin{cases} 0 & \text{if } L < 0 \text{ or } L = 0 \text{ and } h \neq 0 \\ 1 & \text{if } L \geq 0 \text{ and } h = 0 \\ s(L, h \bmod L) & \text{if } L > 0 \text{ and } (h < 0 \text{ or } h > L) \\ s(L, L-h) & \text{if } L > 0 \text{ and } L/2 < h \leq L \end{cases}$$

and

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

otherwise.

Sample values of $s(L, h)$

$L \backslash h$	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

Generating function

$$G_h(X) = \sum_{L \geq 0} s(L,h) X^L \iff s(L,h) = \frac{G_h^{(L)}(0)}{L!}$$

Theorem

$$G_0(X) = \frac{1}{1-X}, \qquad G_1(X) = \frac{X}{(1-X)^2}$$

and for $h \geq 2$,

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

Asymptotic behaviour

$$\begin{aligned} G_h(X) &= \frac{F_h(X)}{(1-X^{h-1})(1-X^h)(1-X^{h+1})} \\ &= \frac{R_h(X)}{(1-X)^3} + \frac{A_h(X)}{1-X^{h-1}} + \frac{B_h(X)}{1-X^h} + \frac{C_h(X)}{1-X^{h+1}} \end{aligned}$$

where $\deg(R_h) < 3$, $\deg(A_h) < h-1$, $\deg(B_h) < h$ and $\deg(C_h) < h+1$.

Corollary

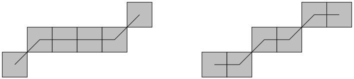
For all $h \geq 2$, there exist $u_0, \ldots, u_{h-2}, v_0, \ldots, v_{h-1}, w_0, \ldots, w_h \in \mathbb{Q}$ such that

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

where

$$\begin{aligned} \alpha &= \frac{\sum_{r=0}^{h-2} s(h-1,r)}{h\,(h^2-1)} = \frac{1}{h\,(h^2-1)} \sum_{i=1}^{h-1} (h-i)\varphi(i) \\ \beta &= \frac{\sum_{r=0}^{h-1} s(h,r) - \sum_{r=0}^{h-2} s(h-1,r)}{h\,(h+1)} = \frac{1}{h\,(h+1)} \sum_{i=1}^{h} \varphi(i) \end{aligned}$$

Symmetric discrete segments



$w = 10001$ $w = 01010$

FIG.: The 2 symmetric segments of length 5 and height 2 and the corresponding encodings.

Symmetric segments are encoded by balanced palindromes.

Questions

1. How many balanced palindromes exist of a given length ?
2. How many balanced palindromes exist of a given length L and height h ?

Question

How many balanced palindromes exist of a given length ?

Theorem (De Luca & de Luca, 2005)

The number $p(L)$ of balanced palindromes of length L is

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

Notations

- Let $\mathbb{P}$ denote the set of finite balanced palindromes over the alphabet $\{0, 1\}$.
- For all $L, h \in \mathbb{N}$ such that $0 \leq h \leq L$,

$$P(L, h) = \{w \mid w \in \mathbb{P},\ |w| = L \text{ and } |w|_1 = h\}$$

and for all $L, h \in \mathbb{Z}$,

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

Recurrence formula

Theorem

$$p(L,h) = \begin{cases} 0 & \textit{if } L < 0 \textit{ or } L = 0 \textit{ and } h \neq 0 \\ 1 & \textit{if } L \geq 0 \textit{ and } (h = 0 \textit{ or } h = L) \\ L \bmod 2 & \textit{if } L \geq 0 \textit{ and } (h = 1 \textit{ or } h = L-1) \\ p(L, h \bmod L) & \textit{if } L > 0 \textit{ and } (h < 0 \textit{ or } h > L) \end{cases}$$

and

$$p(h-1, (L-1) \bmod (h-1)) + p(L-h-1, (L-1) \bmod (L-h-1))$$

otherwise.

Sample values of $p(L, h)$

$L \backslash h$	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

Generating function

$$G_h(X) = \sum_{L \geq 0} p(L,h) X^L \iff p(L,h) = \frac{G_h^{(L)}(0)}{L!}$$

Theorem

$$G_0(X) = \frac{1}{1-X}, \quad G_1(X) = \frac{X}{1-X^2}$$

and for $h \geq 2$,

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

We get for instance :

$$G_2(X) = \frac{X}{(1-X)(1-X^3)} \qquad G_3(X) = \frac{X}{(1-X^2)(1-X^4)}$$

$$G_4(X) = \frac{X}{(1-X)(1-X^5)} \qquad G_5(X) = \frac{X(1+X^2+X^6)}{(1-X^4)(1-X^6)}$$

Asymptotic behaviour

$$G_h(X) = \frac{F_h(X)}{(1 - X^{h-1})(1 - X^{h+1})}$$

If h is even, $G_h(X) = \frac{R_h(X)}{(1-X)^2} + \frac{A_h(X)}{1 - X^{h-1}} + \frac{B_h(X)}{1 - X^{h+1}}$

If h is odd, $G_h(X) = \frac{Q_h(X)}{(1+X)^2} + \frac{R_h(X)}{(1-X)^2} + \frac{A_h(X)}{1 - X^{h-1}} + \frac{B_h(X)}{1 - X^{h+1}}$

where $\deg(Q_h) < 2$, $\deg(R_h) < 2$, $\deg(A_h) < h - 1$ and $\deg(B_h) < h + 1$.

Corollary

If h is even, then there exist $u_0, \ldots, u_{h-2}, v_0, \ldots, v_h \in \mathbb{Q}$ such that

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

If h is odd, then there exist $u_0, \ldots, u_{h-2}, v_0, \ldots, v_h \in \mathbb{Q}$ such that

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

where

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

Perspectives

1. Find a closed formula for the generating functions which do not depend on $s(L, h)$ and $p(L, h)$.
2. Adapt the counting method to $m \times n$ local configurations of discrete planes.

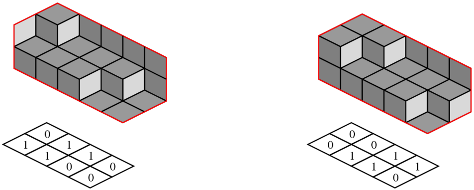


FIG.: $2 \times n$ planar configurations.