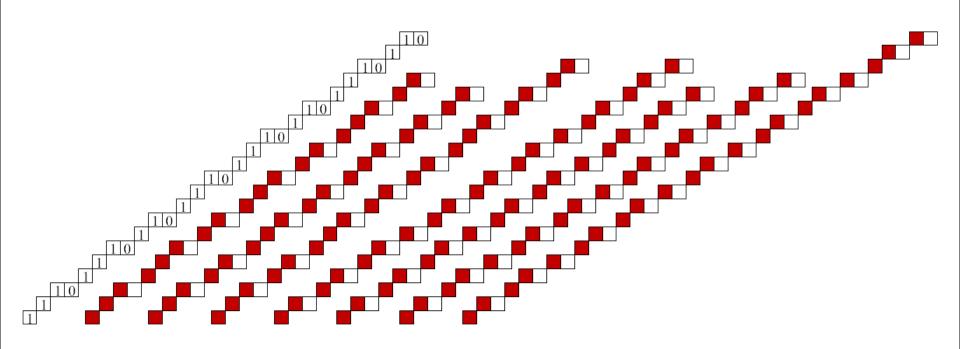


Digital lines, Sturmian words, and continued fractions

Hanna Uscka-Wehlou



25 September 2009



Four short stories:

about continued fractions

about words

about digital lines

about my contribution



1

continued fractions



Continued fractions - notation

$$a = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}} = [0; a_1, a_2, a_3, \dots]$$



Continued fractions - a definition

$$a = [a_0; a_1, a_2, a_3, \dots]$$

$$\alpha_0 = a;$$
 for $n \ge 0$:
$$(a_n) = \lfloor \alpha_n \rfloor, \quad \alpha_{n+1} = \frac{1}{\alpha_n - a_n}$$

$$= \frac{1}{\operatorname{frac}(\alpha_n)}$$



Why do we need continued fractions?

This works for irrational numbers as well

An excellent tool for approximations

They give a better understanding of a number

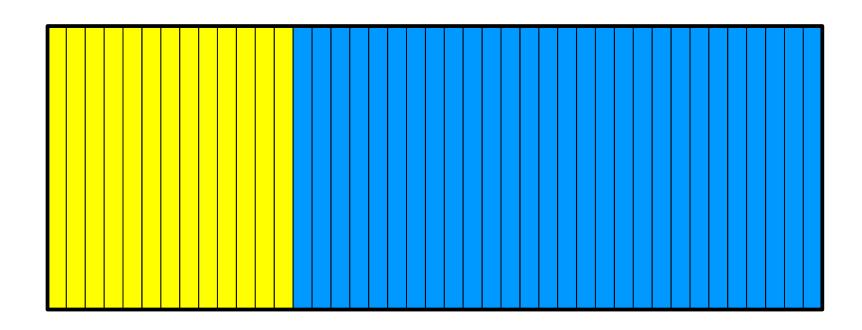
Numerous geometrical interpretations and applications

Solution's techniques in number theory (Pell's equation)



A fraction - an interpretation

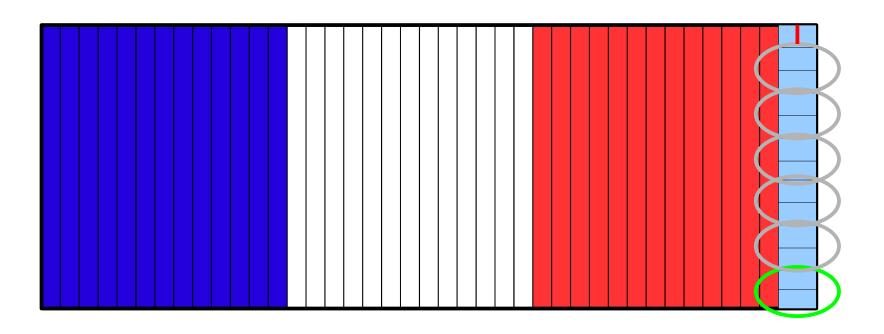
 $\frac{13}{41}$





A continued fraction - an interpretation

$$\frac{13}{41} = \frac{1}{\frac{41}{13}} = \frac{1}{3 + \frac{2}{13}} = \frac{1}{3 + \frac{1}{\frac{13}{2}}} = \frac{1}{3 + \frac{1}{\frac{13}{2}}} = [0; 3, 6, 2].$$





Continued fractions and their relatives

Euclid's algorithm

The Gauss map

The Stern - Brocot tree



2

words

Words - the alphabet

 $A = \{a,b,c,d,e,f,g,h,i,j,k,l,m,n,o,p,q,r,s,t,u,v,w,x,y,z\}$ dziewczynka, vriendschap, dataspel, ciotka, siostrzyczka

$$A = \{ \#, \#, \&, @ \}$$
 $\# \# \# \& @ @ @ \& \& \& \# \# @ \& ...$



Finite words

- A alphabet (a set of symbols)
- A^* the set of finite words over A
- $(A^*, +)$ is a monoid:
 - concatenation (+) is associative (u+v)+w=u+(v+w) 101010+1111=1010101111
 - the empty word is the neutral element
 - $(A^*, +)$ is called the free monoid on the set A.
 - no inverse operation, no commutativity

The set of all right infinite words over {1,2}:

$$\{1,2\}^{\omega}$$

$$w: \mathbf{N}^+ \to \{1, 2\}$$

$$w = w(1)w(2)w(3) \cdots \in \{1, 2\}^{\omega}$$

The run-length encoding operator

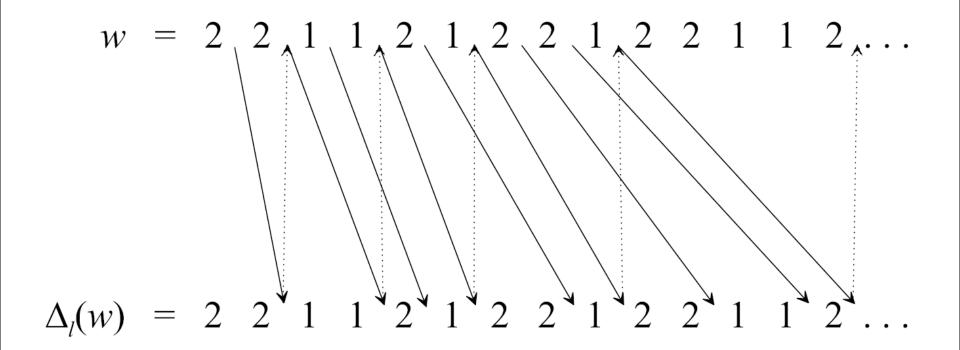
$$\Delta_l: \{1,2\}^\omega \to \mathbf{N}^\omega$$

$$w = \begin{cases} 1^{k_1} 2^{k_2} 1^{k_3} 2^{k_4} \cdots, & \text{if } w \in 1 \cdot \{1, 2\}^{\omega} \\ 2^{k_1} 1^{k_2} 2^{k_3} 1^{k_4} \cdots, & \text{if } w \in 2 \cdot \{1, 2\}^{\omega} \end{cases}$$
$$\Delta_l(w) = k_1 k_2 k_3 \cdots$$

The run-length encoding operator - an example:

$$w = 11111212222222112222222221...$$

 $\Delta_{l}(w) = 5,1,1,7,2,3,1,1,1,9,...$





The word w is called a factor of a word u if there exist words x, y such that u=x+w+y.

1222 is a factor of 000(1222)11113213110101001

10101 is a factor of (101010101010101

ABCDA is a factor of CBBBDACADBCAABCDA

10101 is a factor of 10101



Sturmian words are infinite words which have exactly m+1 different factors of length m for every natural m.



Sturmian words are infinite words which have exactly m+1 different factors of length m for every natural m.

m=1 — two letters (binary words)



Sturmian words are infinite words which have exactly m+1 different factors of length m for every natural m.

m=1 — two letters (binary words)



Sturmian words are infinite words which have exactly m+1 different factors of length m for every natural m.

m=1 — two letters (binary words)

m=4

1010, 0101, 0010, 1001, 0100.



Sturmian words are infinite words which have exactly m+1 different factors of length m for every natural m.

m=1 — two letters (binary words)

m=4

1010, 0101, 0010, 1001, 0100.



Sturmian words are infinite words which have exactly m+1 different factors of length m for every natural m.

m=1 — two letters (binary words)

m=4

1010, 0101, 0010, 1001, 0100.



Balanced words (binary)

n - the length of the word

m - any positive natural number less than n

Each m-letter long factor of this word can contain either k or k+1 1's

An example:

$$n = 41$$

$$m = 16$$

10101001010010100101001010010100100100

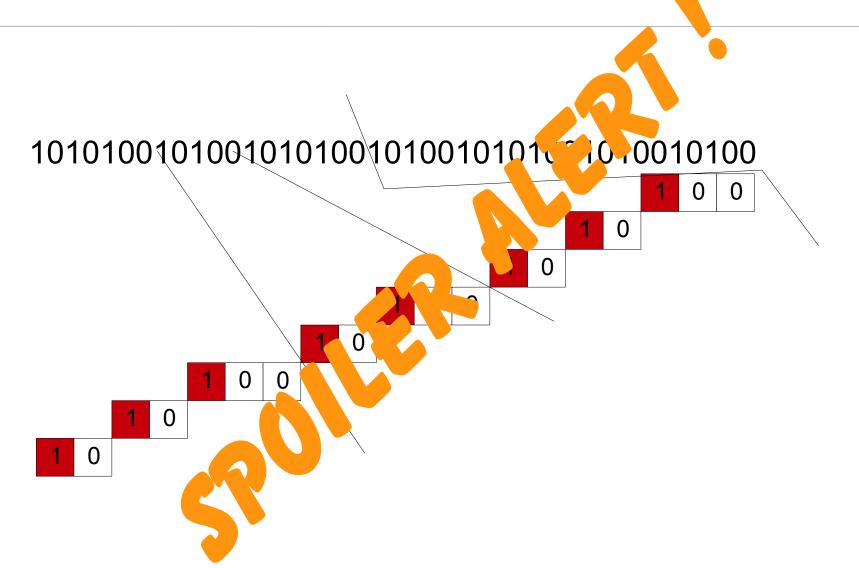
7

$$k = 6$$

6



Balanced words give straight lines





Upper and lower mechanical, characteristic words

$$s'(a), s(a): \mathbf{N} \to \{0, 1\}$$

$$\forall n \in \mathbf{N} \quad s'_n(a) = \lceil a(n+1) \rceil - \lceil an \rceil,$$

$$s_n(a) = \lfloor a(n+1) \rfloor - \lfloor an \rfloor$$

$$c(a): \mathbf{N}^+ \to \{0, 1\}$$

$$\forall n \in \mathbf{N}^+ \quad c_n(a) = \lfloor a(n+1) \rfloor - \lfloor an \rfloor$$

Sturmian words: different characterizations

Theorem Let s be an infinite word. The following are equivalent:

- \bullet s is Sturmian;
- s is balanced and aperiodic;
- s is irrational (lower or upper) mechanical.



Combinatorics on words

Palindromes

kajak, Hannah, wow, SOS, 123454321, ?*#@#*?

Square-free words

adbcbabcbdabdcbabcdbabdcb ...

Overlap-free words

abbabaabbaababaab ... (the Thue-Morse word)

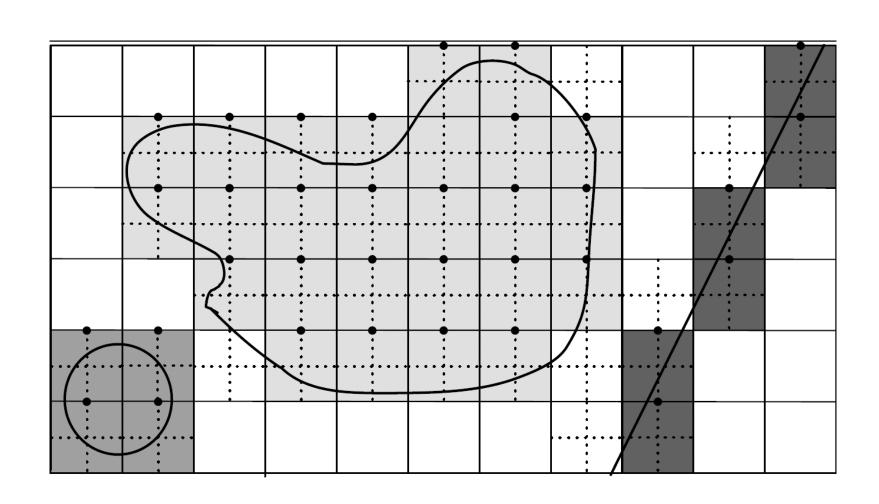


3

digital lines



Digital geometry - R'-digitization



Digital geometry - straight lines

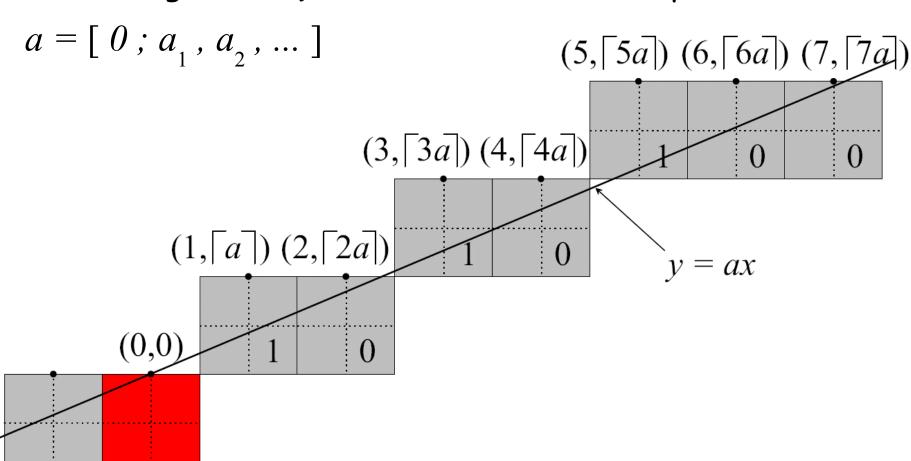
The arithmetical expression of the R'-digitization of the line y = ax for irrational positive a less than 1:

$$D_{R'}(y = ax) = \{(k, \lceil ak \rceil); k \in \mathbf{Z}\}$$



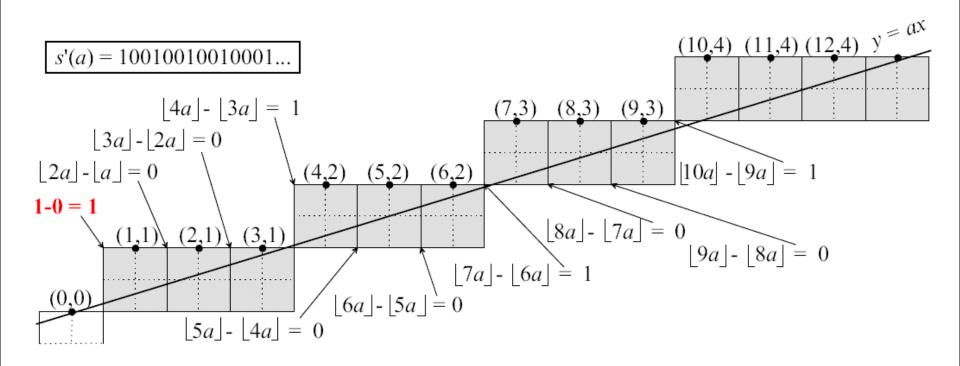
Digital geometry - straight lines

The R'-digital line y = ax with irrational slope

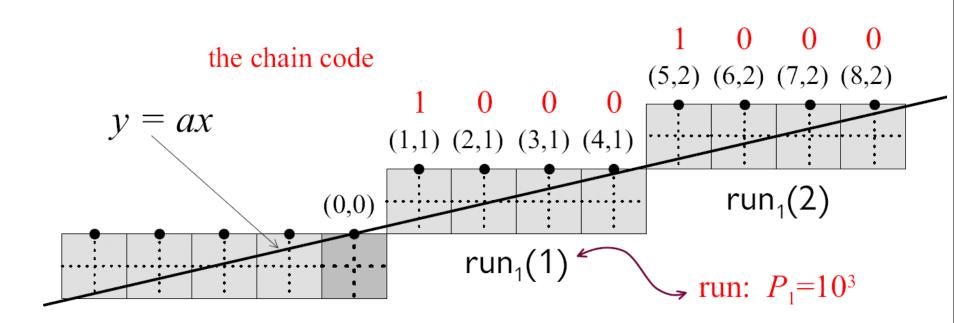


Digital geometry - straight lines and mechanical words

The R'-digital line y = ax with slope $a = [0; a_1, a_2, ...]$ and the corresponding upper mechanical word s'(a):

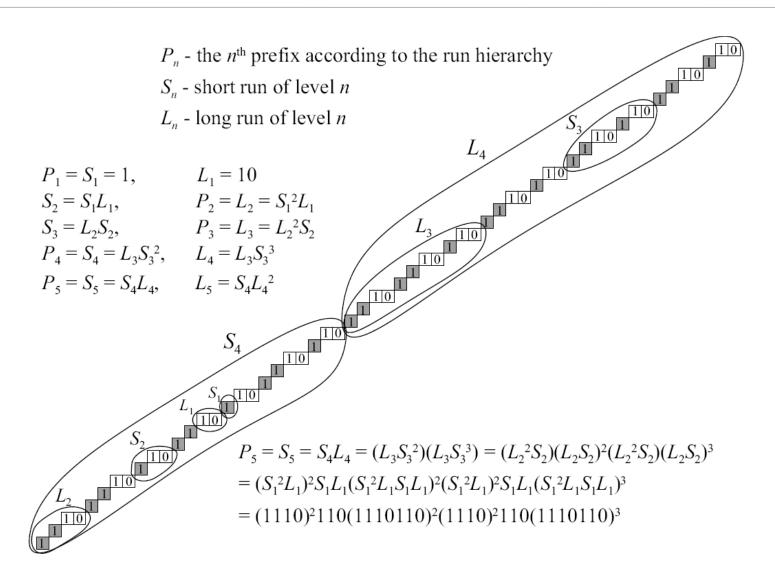


Digital geometry - the concept of run





Digital geometry - the concept of run





Hierarchy of runs - runs on level k+1

$$L_k S_k^m$$

$$L_k S_k^m \qquad S_k^m L_k \qquad L_k^m S_k \qquad S_k L_k^m$$

$$L_k^m S_k$$

$$S_k L_k^m$$



Three questions. About:

the run length on level k+1

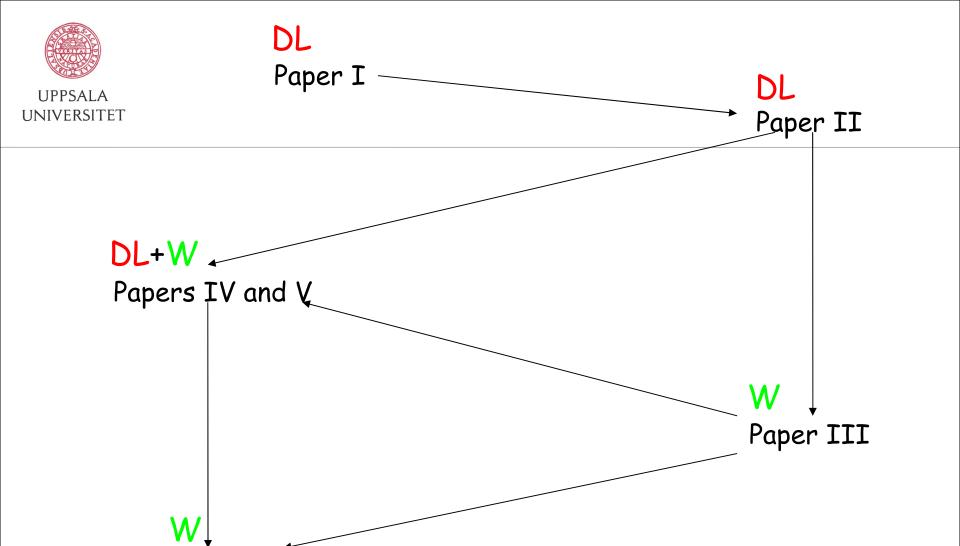
the main run on level k

the first run on level k



4

my contribution



Paper VI



DL

Paper I

Digitization parameters:

DL

Paper II

Paper III

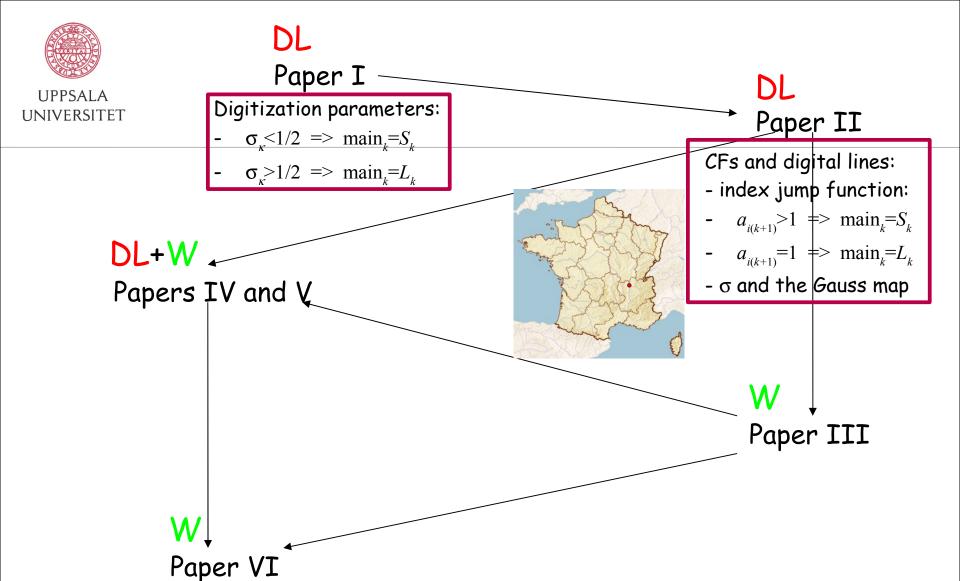
- $\sigma_{\kappa} < 1/2 => \min_{k} = S_{k}$
- $\sigma_{\kappa} > 1/2 => \min_{k} = L_{k}$

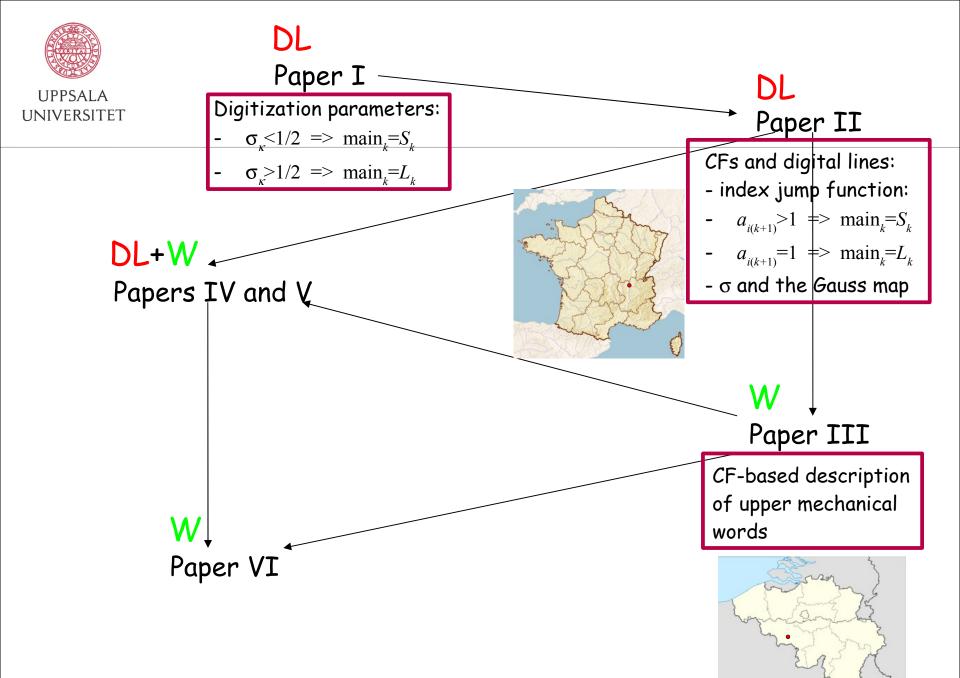
DL+W .

Papers IV and V

W

Paper VI







DL

Paper I

Digitization parameters:

- $\sigma_{\kappa} < 1/2 => \min_{k} = S_{k}$
- $\sigma_{\kappa} > 1/2 => \min_{k} = L_{k}$

DL+W .

Papers IV and V

Two equivalence relations:

- by run length
- by the index jump function

W↓

Paper VI

DL Paper II

CFs and digital lines:

- index jump function:
- $a_{i(k+1)} > 1 \neq \min_{k} S_k$
- $a_{i(k+1)}=1 \neq \min_{k}=L_k$
- σ and the Gauss map

W

Paper III

CF-based description of upper mechanical words





DL

Paper I

Digitization parameters:

- $\sigma_{\kappa} < 1/2 => \min_{k} = S_{k}$
- $\sigma_{\kappa} > 1/2 => \min_{k} = L_{k}$

DL Paper II

CFs and digital lines:

- index jump function:
- $a_{i(k+1)} > 1 \neq \min_{k} S_k$
- $a_{i(k+1)}=1 \neq \min_{k}=L_{k}$
- σ and the Gauss map

DL+W .

Papers IV and V

Two equivalence relations:

- by run length
- by the index jump function

W

Paper III

CF-based description of upper mechanical words



Paper VI

- the constructional word
- balanced construction
- fixed point theorem





What I have done

A new CF-description (essential 1's, run hierarchy)

Two equivalence relations on the set of slopes

A new fixed point theorem for words



The index jump function

The index jump function

$$a = [0; a_1, a_2, a_3, a_4, a_5, a_6, a_7, \dots]$$

 $i_a : \mathbf{N}^+ \to \mathbf{N}^+$

$$i_a(1) = 1$$
, $i_a(2) = 2$, for $n \ge 2$:
 $i_a(n+1) = i_a(n) + 1 + \delta_1(a_{i_a(n)})$



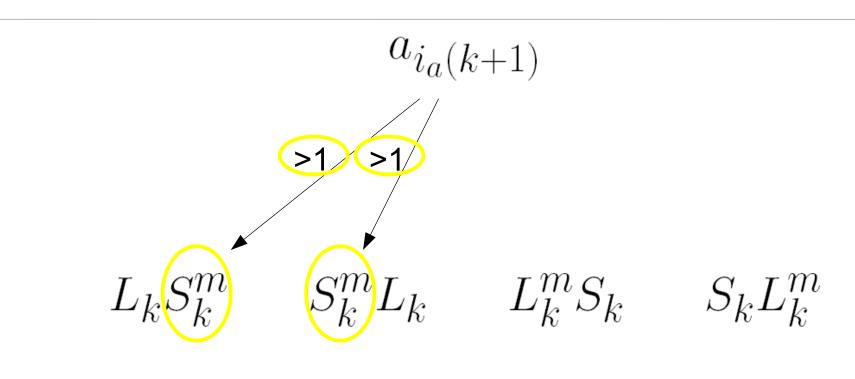
How $i_a(k+1)$ and $a_{i_a(k+1)}$ describe the form of run_{k+1}

$$a_{i_a(k+1)}$$

$$i_a(k+1)$$

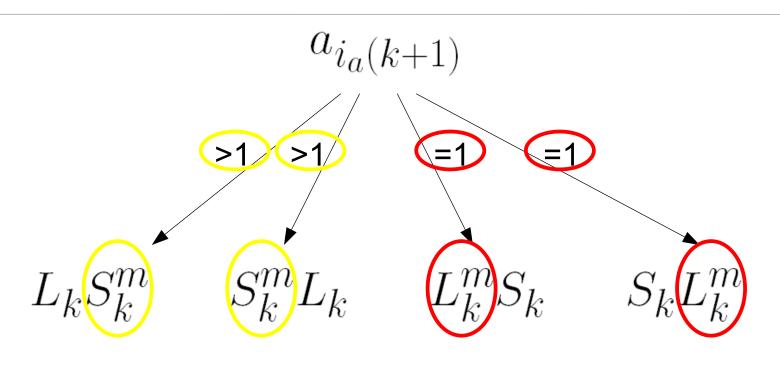
 $L_k S_k^m \qquad S_k^m L_k \qquad L_k^m S_k \qquad S_k L_k^m$





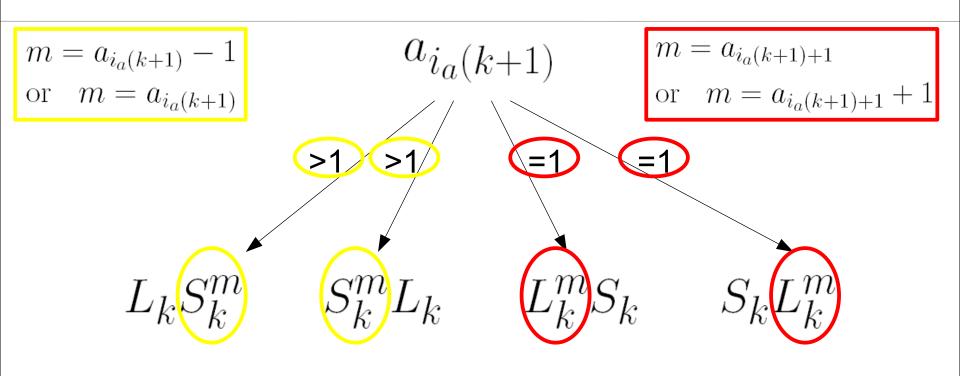
$$i_a(k+1)$$





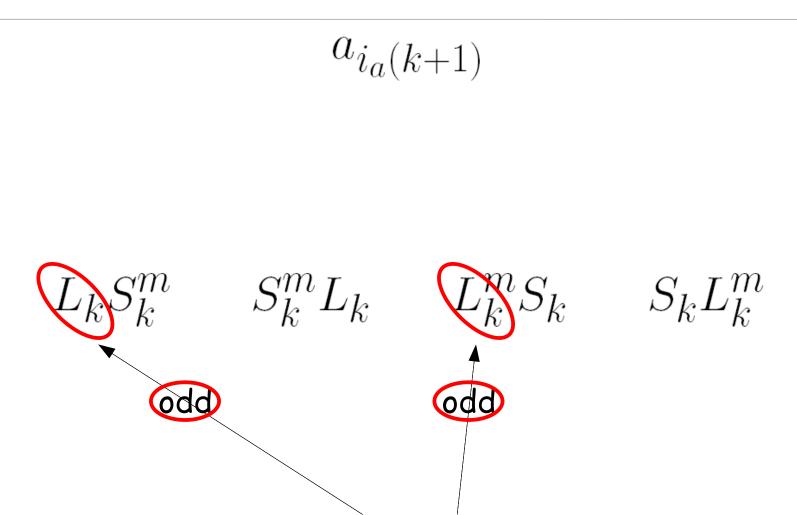
$$i_a(k+1)$$





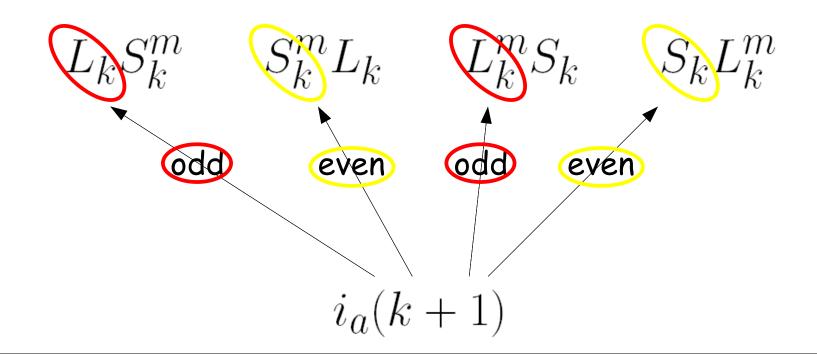
$$i_a(k+1)$$







$$a_{i_a(k+1)}$$



The index jump function: how it describes the runs



Digitization levels

$a = [0; \stackrel{b_1}{1}, \stackrel{b_2}{a_2}, \stackrel{b_3}{1}, \stackrel{b_4}{1}, \stackrel{b_5}{a_5}, \stackrel{b_6}{1}, \stackrel{b_7}{1}, \stackrel{b_8}{a_9}, \stackrel{b_7}{1}, \stackrel{b_8}{a_{11}}, \stackrel{b_9}{a_{12}}, \stackrel{b_{10}}{1}, \stackrel{b_{11}}{1}, \stackrel{b_{12}}{a_{16}}, \stackrel{b_{12}}{a_{17}}, \dots]$ $(i_a(k))_{k \in \mathbb{N}^+} = (1, 2, 3, 5, 6, 8, 9, 10, 12, 13, 15, 15, 17, \dots)$	Level:	1	2	3	4	5	6	7	8	9	10	11	12	
								·)



Short run length: the CF elements

Level: 1	2	3	4	5	6	7	8	9	10	11	12	
$a = [0; \stackrel{b_1}{1}]$ $\downarrow (i_a(k))_{k \in \mathbb{N}^+} = (1)$	\downarrow	$\underbrace{\overbrace{1, \ 1}^{b_3}}_{3,},$				\downarrow	$\underbrace{1, a_{11}}_{b_8},$ \downarrow $10,$	$ \begin{array}{c} a_{12}^{b_9}, \\ \downarrow \\ 12, \end{array} $	\downarrow	$\underbrace{\frac{1}{1}, a_{16}}^{b_{11}},$ \downarrow $15,$	$a_{17}^{b_{12}},\ \downarrow \ 17,$]
$ S_k = b_k$												



The most frequent run: essential 1's

Level: 1	2	3	4	5	6	7	8	9	10	11	12	
$a = \begin{bmatrix} 0; & 1 \\ 1 & \downarrow \\ (i_a(k))_{k \in \mathbf{N}^+} = & (& 1 \end{bmatrix}$	$\begin{pmatrix} b_2 \\ a_2 \\ 2 \end{pmatrix}$	11	$\overset{b_4}{a_5},\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	$\underbrace{\overbrace{1}^{b_5}, 1}_{c, b_5},$	$\overset{b_6}{a_8}, \\ \downarrow \\ 8, \\ \end{matrix}$	$\overset{b_7}{\overset{a_9}{{{{{{}{}{}{$	/ 1	/1	/1	$ \underbrace{1, a_{16}}_{b_{11}}, \\ \downarrow \\ 15, $	$\begin{array}{c} b_{12} \\ a_{17}, \\ \downarrow \\ 17, \end{array}$]
$egin{array}{cccccccccccccccccccccccccccccccccccc$	$egin{pmatrix} L_2 \end{pmatrix}$	S_3	L_4	S_5	S_6	L_7	S_8	L_{9}	L_{10}	S_{11}	?	
$a_{i_{d}(k+1)} > 1 \qquad a_{i_{d}(k+1)} = 1$		3	,	3	Ü	,	Ü		10			



The first run: parity of the function

Level :	1	2	3	4	5	6	7	8	9	10	11	12	
	\downarrow	$\overset{b_2}{a_2}, \ \downarrow \ 2,$	1	ı	1	1	1	$\underbrace{1, a_{11}}_{b_8},$ \downarrow $10,$	1	1	\downarrow	\downarrow]
$(i_a(k))_{k \in \mathbb{N}^+} =$	(1,	2,	3,	5,	6 ,	8,	9,	10,	12,	$\overset{\downarrow}{\overset{1}{3}},$	15,	17,)
first _k $i_a(k+1)$ even $i_a(k+1)$		$\stackrel{ullet}{L_2}$	L_3	S_4	S_{5}	$\stackrel{ullet}{L}_6$	S_7	S_8	$\stackrel{ullet}{L_9}$	L_{10}	L_{11}	?	

Two equivalence relations on the set of slopes

1. based on run length on all levels for s'(a):

$$a \in [(b_1, b_2, b_3, \dots)]_{\sim_{\text{len}}} \Leftrightarrow$$

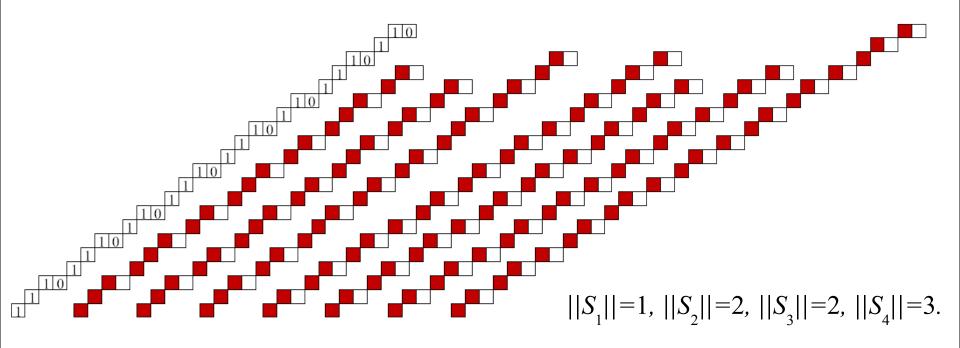
$$\forall k \in \mathbf{N}^+ ||S_k|| = b_k$$

2. based on run construction on all levels for s'(a):

$$a \sim_{\text{con}} a' \Leftrightarrow i_a \equiv i_{a'}$$

Quantitative equivalence relation (run length)

Defined by run lengths (their cardinality)



All lines from the same class have the same run lengths on all digitization levels.

Quantitative equivalence relation (run length)

The least element of the class:

$$\min\{a \in]0, 1[\setminus \mathbf{Q}; \ a \in [(b_n)_{n \in \mathbf{N}^+}]_{\sim_{\text{len}}} \} = [0; b_1, \overline{1, b_n - 1}]_{n=2}^{\infty}.$$

The largest element of the class:

$$\max\{a \in]0, 1[\ \setminus \mathbf{Q}; \ a \in [(b_n)_{n \in \mathbf{N}^+}]_{\sim_{\text{len}}}\} = [0; b_1, b_2, \overline{1, b_n - 1}]_{n=3}^{\infty}.$$

Qualitative equivalence relation (run construction)

Defined by the places of essential 1's

Equivalently defined by the index jump function

All lines from the same class have the same construction in terms of long and short runs on all digitization levels.

The least element in each class is 0.



Qualitative equivalence relation (run construction)

Largest elements of the classes:

$$\forall n \in \mathbf{N}^{+} [(\forall k \in [1, n-1]_{\mathbf{Z}} \ s_{k} = 2k)$$

$$\land (s_{n} > 2n \ \lor \ |J| = n-1)]$$

$$\Rightarrow \sup\{a; \ a \in [(s_{j})_{j \in J}]_{\sim_{\text{con}}}\} = \frac{F_{2n-1}}{F_{2n}},$$

where $(F_n)_{n \in \mathbb{N}^+}$ is the **Fibonacci** sequence and

 $(s_j)_{j\in J}$ is the sequence of the places of essential 1's.



Fixed point theorem: the constructional word

The constructional word $\gamma(a) \in \{0,1\}^{\omega}$

Let
$$a = [0; a_1, a_2, \ldots]$$
. For $n \in \mathbf{N}^+$:

$$\gamma_n(a) = i_a(n+2) - i_a(n+1) - 1$$

$$\gamma_n(a) = \delta_1(a_{i_a(n+1)})$$

$$\gamma_n(a) = \begin{cases} 0, & S_n \text{ is the most frequent} \\ & \text{run on level } n \text{ for } s'(a) \\ 1, & L_n \text{ is the most frequent} \\ & \text{run on level } n \text{ for } s'(a). \end{cases}$$



Fixed point theorem: the constructional word

The constructional word $\gamma(a) \in \{0,1\}^{\omega}$

Let
$$a = [0; a_1, a_2, \ldots]$$
. For $n \in \mathbf{N}^+$:

$$\gamma_n(a) = i_a(n+2) - i_a(n+1) - 1$$

$$\gamma_n(a) = \delta_1(a_{i_a(n+1)})$$

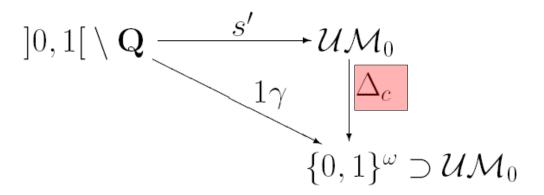
$$\gamma_n(a) = \begin{cases} 0, & S_n \text{ is the most frequent} \\ & \text{run on level } n \text{ for } s'(a) \\ 1, & L_n \text{ is the most frequent} \\ & \text{run on level } n \text{ for } s'(a). \end{cases}$$



Fixed point theorem: the run-construction encoding operator

Definition The run-construction encoding operator

 $\Delta_c: \mathcal{UM}_0 \longrightarrow \{0,1\}^{\omega} \text{ is defined as } \Delta_c = (1\gamma) \circ (s')^{-1}.$



where \mathcal{UM}_0 denotes the set of all upper mechanical words with irrational slope 0 < a < 1 and with intercept 0.



Balanced construction

Let $a \in]0,1[\setminus \mathbf{Q}]$. The word s'(a)=1c(a) has balanced construction if

$$\exists \ \alpha \in \mathbf{R} \quad \gamma(a) = c(\alpha)$$

Sturmian-balanced construction if

$$\exists \alpha \in]0,1[\setminus \mathbf{Q} \quad \gamma(a) = c(\alpha)$$

<u>self-balanced construction</u>

$$1\gamma(a) = \Delta_c(1c(a)) = 1c(a)$$

Fixed point theorem

A fixed-point theorem: exactly 1 fixed point in each equivalence class

Let
$$(b_n)_{n \in \mathbb{N}^+}$$
 be such that $b_1 \in \mathbb{N}^+$
and $b_n \in \mathbb{N}^+ \setminus \{1\}$ for all $n \geq 2$. Then
$$\exists_{a \in]0,1[\setminus \mathbf{Q}}^1$$
 $a \in [(b_n)_{n \in \mathbb{N}^+}]_{\sim_{\text{len}}} \land s'(a) = \Delta_c(s'(a)).$

Fixed point theorem

A fixed-point theorem: exactly 1 fixed point in each equivalence class

Let
$$(b_n)_{n \in \mathbb{N}^+}$$
 be such that $b_1 \in \mathbb{N}^+$ and $b_n \in \mathbb{N}^+ \setminus \{1\}$ for all $n \geq 2$. Then
$$\exists_{a \in]0,1[\setminus \mathbf{Q}}^1$$
 $(a \in [(b_n)_{n \in \mathbb{N}^+}]_{\sim_{\text{len}}}) \land s'(a) = \Delta_c(s'(a)).$

Fixed point theorem

A fixed-point theorem: exactly 1 fixed point in each equivalence class

Let
$$(b_n)_{n \in \mathbb{N}^+}$$
 be such that $b_1 \in \mathbb{N}^+$ and $b_n \in \mathbb{N}^+ \setminus \{1\}$ for all $n \geq 2$. Then
$$\exists_{a \in]0,1[\setminus \mathbf{Q}}^1$$

$$a \in [(b_n)_{n \in \mathbb{N}^+}]_{\sim_{\text{len}}} \land \underline{s'(a)} = \Delta_c(s'(a)).$$



What to do with it?

Digital geometry - operations defined on classes?

Combinatorics on words - new classes of words

Iterations of the run-construction encoding operator

What can one say about the fixed points?

Data compression?



Dziękuję za uwagę

Tack för er uppmärksamhet

Je vous remercie de votre attention

Thank you for your attention

Dank u voor uw aandacht