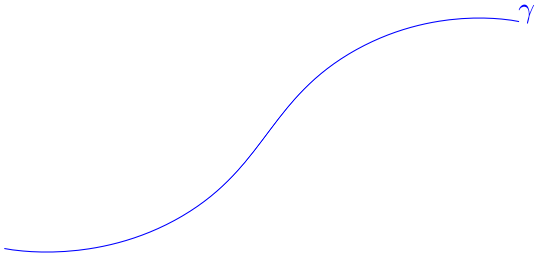
Another definition for digital tangents

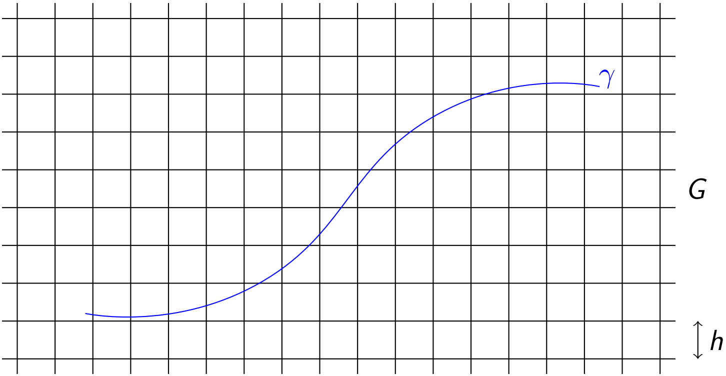
Thierry Monteil

DGCI 2011

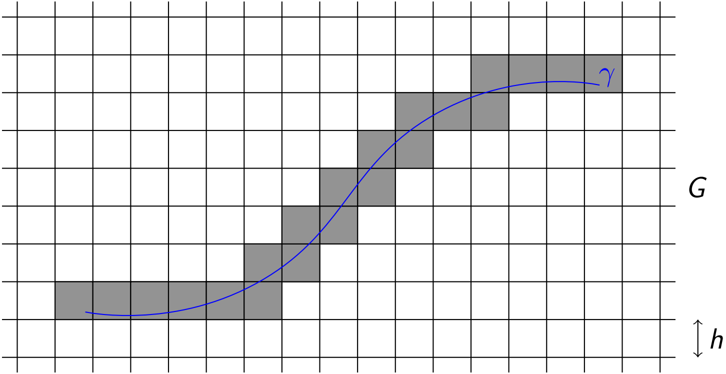
Coding of a smooth curve



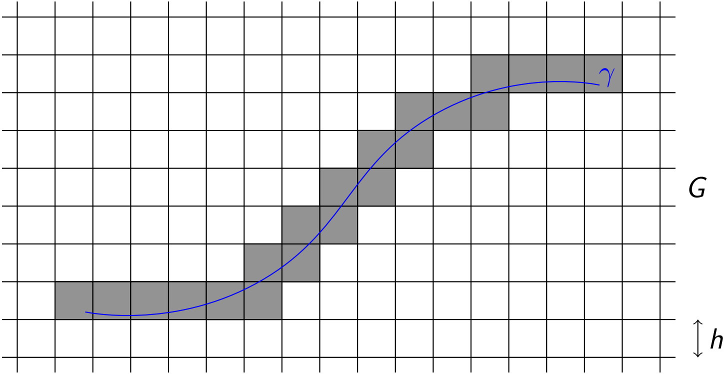
Coding of a smooth curve



Coding of a smooth curve

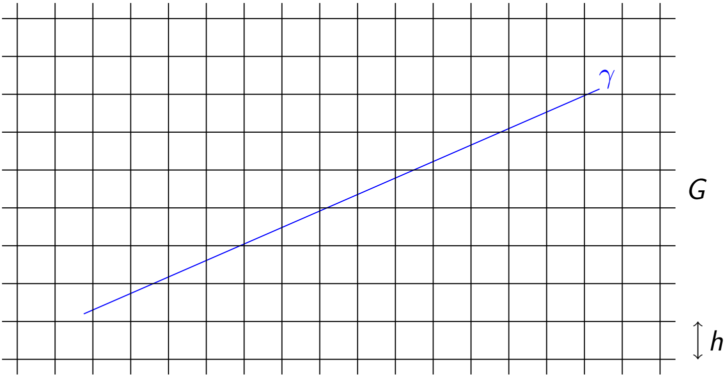


Coding of a smooth curve



$$F(\gamma, G) = 0000010101010101001000$$

The case of digital straight segments



$$w = F(\gamma, G) = 00100100010010010001$$

The case of digital straight segments

There are three types of characterisations of the codings of digital straight segments, corresponding to three properties of the lines:

- *Lines are the curves of constant slope*
  A word w is 1-*balanced* if
  $\forall u, v \in Fact(w) \quad |u| = |v| \Rightarrow ||u|_1 - |v|_1| \leq 1$

The case of digital straight segments

There are three types of characterisations of the codings of digital straight segments, corresponding to three properties of the lines:

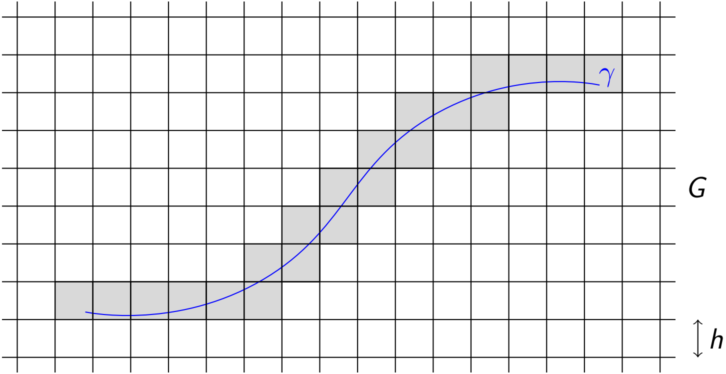
- *Lines are the curves of constant slope*
  A word w is 1-*balanced* if
  $\forall u, v \in Fact(w) \quad |u| = |v| \Rightarrow ||u|_1 - |v|_1| \leq 1$
- *The set of lines is stable under linear maps*
  Multi-scale structure (desubstitution, continued fractions):
  $w = 0010010001001001001 \mapsto 11011101 \mapsto 00 \mapsto \varepsilon$

The case of digital straight segments

There are three types of characterisations of the codings of digital straight segments, corresponding to three properties of the lines:

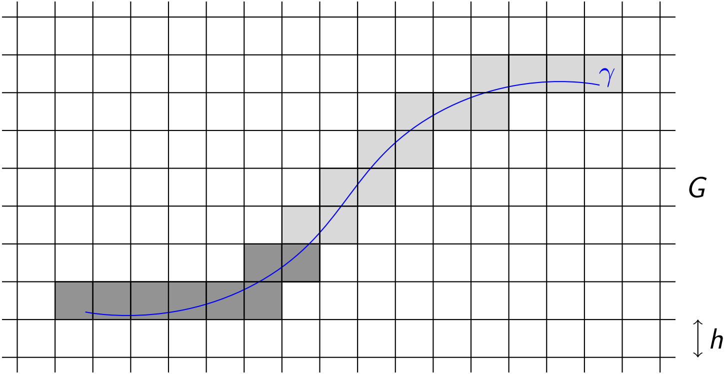
- *Lines are the curves of constant slope*
  A word w is 1-*balanced* if
  $\forall u, v \in Fact(w) \quad |u| = |v| \Rightarrow ||u|_1 - |v|_1| \leq 1$
- *The set of lines is stable under linear maps*
  Multi-scale structure (desubstitution, continued fractions):
  $w = 0010010001001001000 1 \mapsto 11011101 \mapsto 00 \mapsto \varepsilon$
- *Lines are the most predictible curves*
  A word w is a *Sturmian factor* if
  w is a factor of an infinite word with complexity $n + 1$

Segmentation



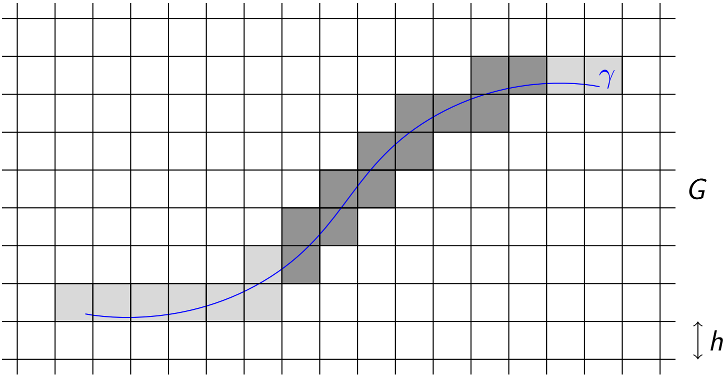
$$F(\gamma, G) = 0000010.10101010010.00$$

Segmentation



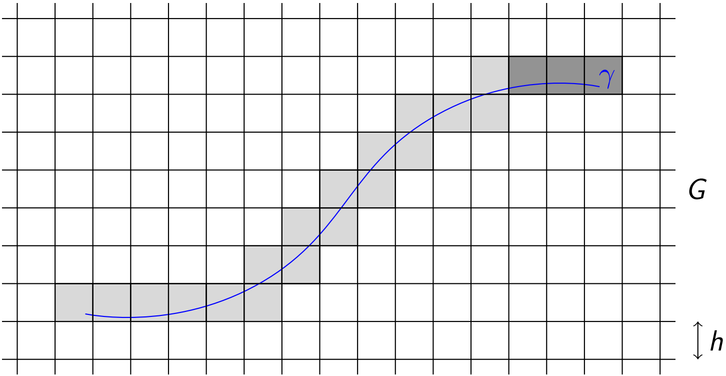
$$F(\gamma, G) = \color{blue}{0000010}\color{black}{.10101010010.00}$$

Segmentation



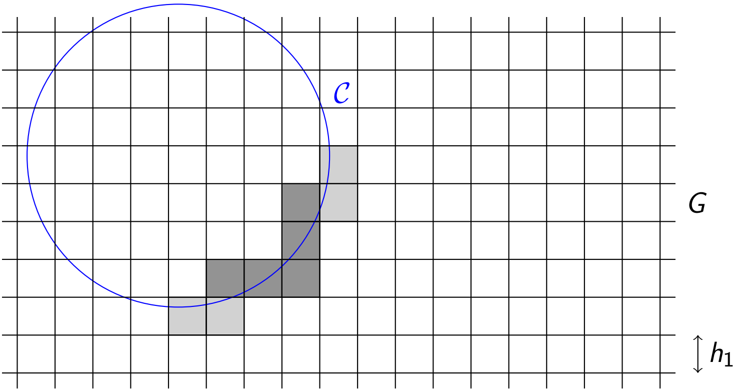
$$F(\gamma, G) = 0000010.10101010010.00$$

Segmentation



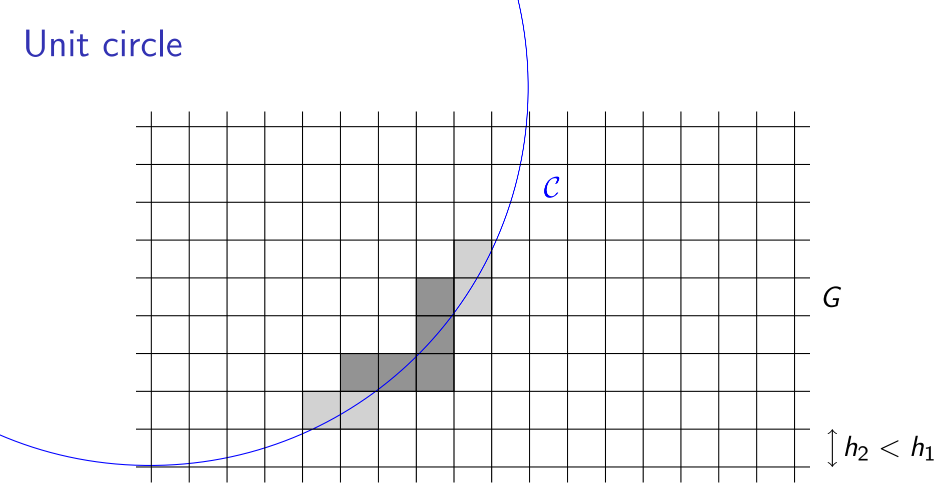
$$F(\gamma, G) = 0000010.10101010010.00$$

Unit circle



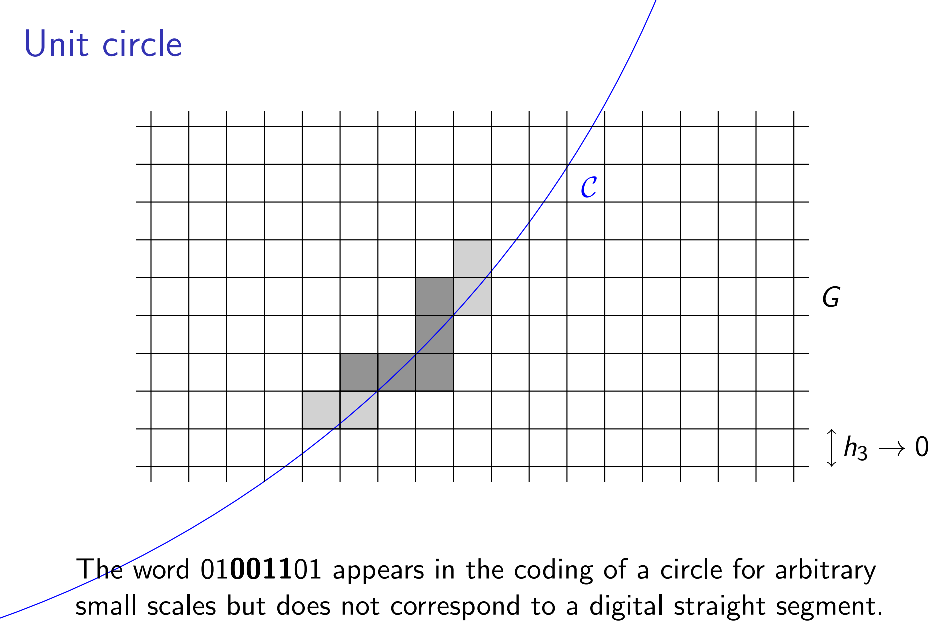
The word 01**0011**01 appears in the coding of a circle for arbitrary small scales but does not correspond to a digital straight segment.

Unit circle



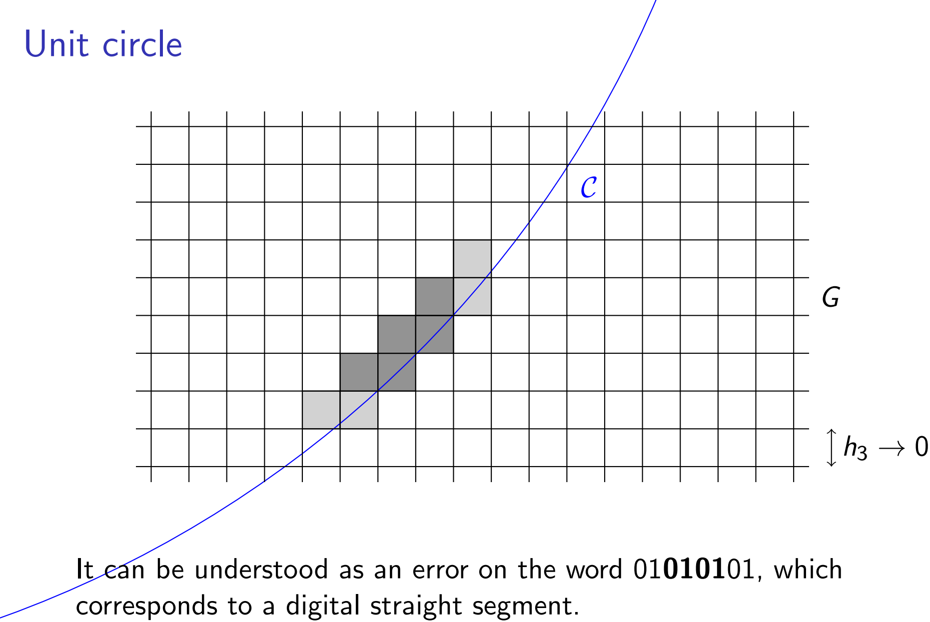
The word 01**0011**01 appears in the coding of a circle for arbitrary small scales but does not correspond to a digital straight segment.

Unit circle



The word 01**0011**01 appears in the coding of a circle for arbitrary small scales but does not correspond to a digital straight segment.

Unit circle



It can be understood as an error on the word 01**0101**01, which corresponds to a digital straight segment.

Tangent words

The aim of this talk is to introduce and describe those words that survive in the coding of a smooth curve when the mesh of th grid goes to 0.

If γ is a smooth curve, we define the tangent words to γ as

$$T(\gamma) = \limsup_{mesh(G)\to 0} L(\gamma, G) = \bigcap_{\varepsilon>0} \bigcup_{mesh(G)\leq\varepsilon} L(\gamma, G),$$

where $L(G, \gamma)$ denotes the set of words appearing in the coding of γ by the grid G.

We denote by T the set of all tangent words obtained for all smooth curves and call them *tangent words*.

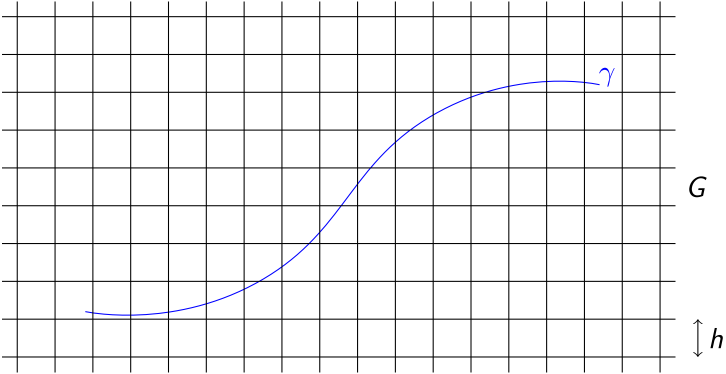
First examples

Any coding of a digital straight segment is tangent.

We just saw that the word 0011 is tangent.

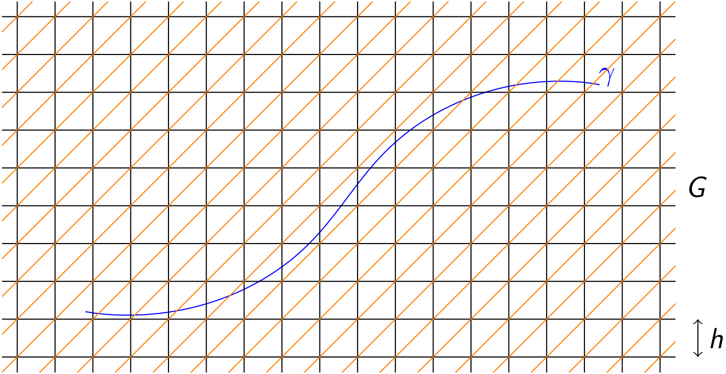
The word 00011 is not tangent.

Combinatorial characterisation: recoding



$w = 000001010101001000$

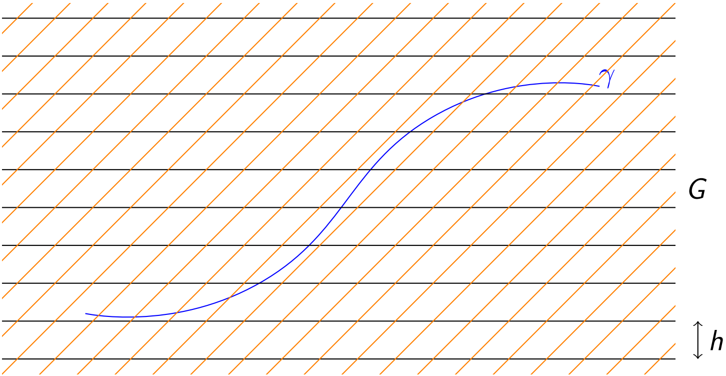
Combinatorial characterisation: recoding



$w = 00000101010101001000$

$0{\color{orange}2}0{\color{orange}2}0{\color{orange}2}0{\color{orange}2}0101010101{\color{orange}2}010{\color{orange}2}0{\color{orange}2}0$

Combinatorial characterisation: recoding

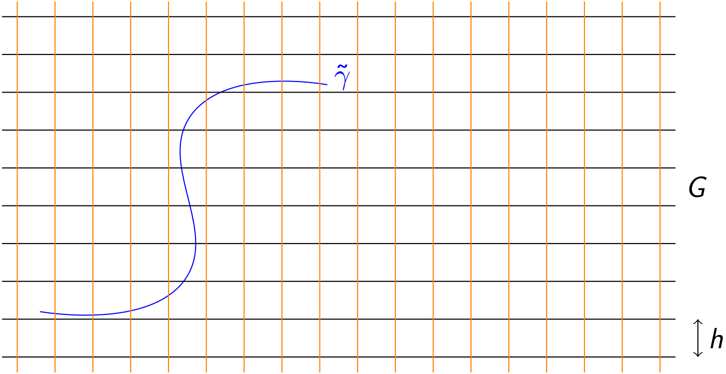


$w = 00000101010101001000$

02020202010101010102010202020

2222111112122

Combinatorial characterisation: recoding



$w = 000001010101010001000$

0202020201010101010201020200

2222111112122

$\tilde{w} = 0000111110100$

We removed one letter to each run of the non-isolated letter.

Combinatorial characterisation: multi-scale structure

A word w is a coding of a curve γ if, and only if, $\tilde{w}$ is a coding of $\tilde{\gamma} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \circ \gamma$ (resp. $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \circ \gamma$ if 0 is isolated).

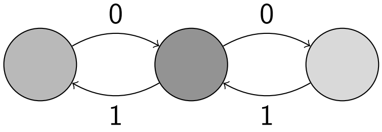
Hence, a word w is tangent if, and only if, $\tilde{w}$ is tangent.

Since $\tilde{w}$ is shorter than w, we can repeat the process until we are stuck: we get a word $d(w)$.

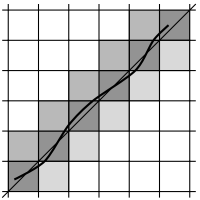
- $d(w)$ is empty if, and only if, w is a coding of a digital straight segment.
- Otherwise, $d(w)$ contains 00 and 11: we have to characterize tangent words whose slope is 1.

Combinatorial characterisation: slope 1

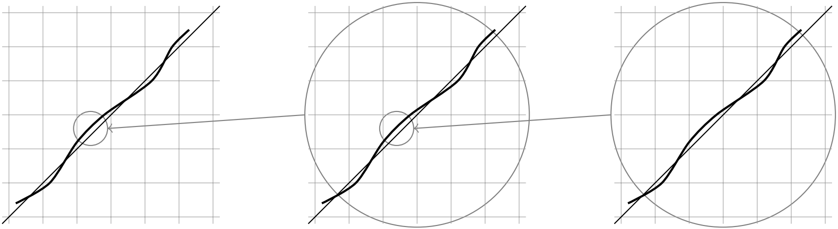
A word w is tangent if, and only if, $d(w)$ is recognised by the following automaton with three states, which are all considered as initial and accepting:



For example, the word 0110100110 (which is not balanced) is diagonal:



Slope 1: how to produce such oscillations ?



An example

$$100100010010010010001001000100 = w$$

An example

100100010010010010001001000100 = w

1~~00~~1~~00~~01~~00~~1~~00~~1~~00~~1~~00~~01~~00~~1~~00~~01~~00~~

An example

$1001000100100100010010001 00 = w$

1~~00~~1~~00~~01~~00~~1~~00~~1~~00~~1~~00~~01~~00~~1~~00~~01~~00~~

110111101101

An example

$100100010010010010001001000100 = w$

100100010010010010001001000100

110111101101

110111101101

An example

$100100010010010001001000100 = w$

$100100010010010001001000100$

110111101101

110111101101

$01100 = d(w)$

An example

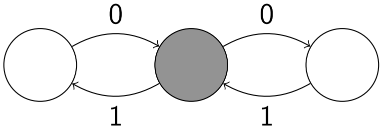
$100100010010010010001001000100 = w$

$100100010010010010001001000100$

110111101101

110111101101

$01100 = d(w)$



An example

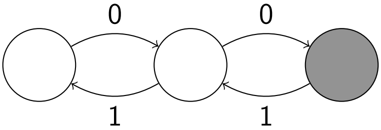
$100100010010010010001001000100 = w$

100100010010010010001001000100

110111101101

110111101101

$01100 = d(w)$



An example

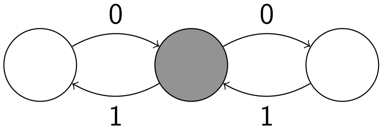
$100100010010010010001001000100 = w$

100100010010010010001001000100

110111101101

110111101101

$01100 = d(w)$



An example

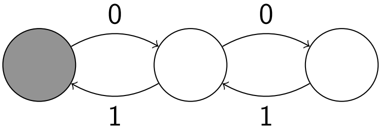
$100100010010010010001001000100 = w$

$100100010010010010001001000100$

110111101101

110111101101

$01100 = d(w)$



An example

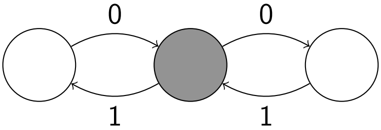
$1001000100100100010010001 00 = w$

10010001001001000100100100 0 100

110111101101

110111101101

$01100 = d(w)$



An example

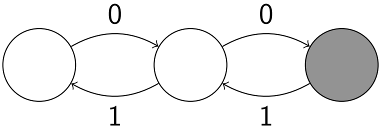
1001000100100100010010001000100 $= w$

1001000100100100010010001000100

110111101101

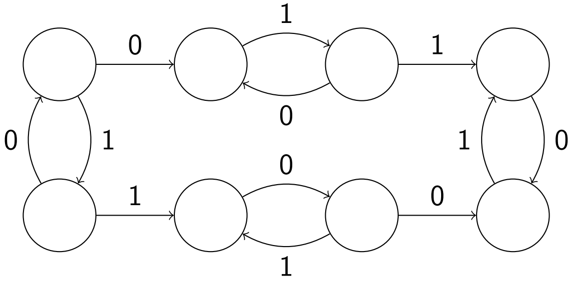
110111101101

01100 $= d(w)$



No bad vibration: analytic curves

A word w is a tangent word of an analytic curve if, and only if, $d(w)$ is recognised by the following automaton with eight states, which are all considered as initial and accepting:



The tangent analytic curves also correspond to the tangent words of the smooth curves with nowhere zero curvature,
For example, 001100 is tangent but not tangent analytic.

Balance and complexity

Each class of words is strictly included in the next one:

- 1-balanced words (digital straight segments)
- tangent analytic words
- tangent (smooth) words
- 2-balanced words

For the first two classes, the complexity is cubical whereas for the last two classes, the complexity is exponential.

Recall:

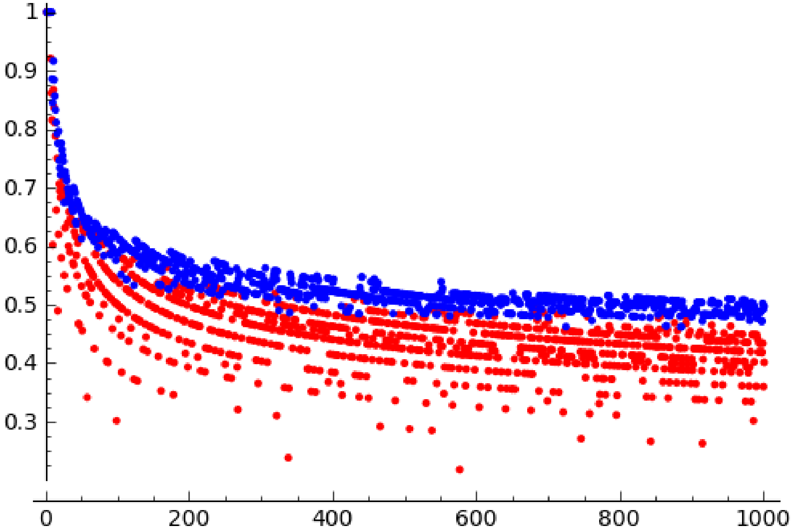
- a word is k-balanced if $\forall u, v \in Fact(w) \quad |u| = |v| \Rightarrow ||u|_1 - |v|_1| \leq k$
- the complexity of a set of words maps each integer n to the number of words of length n.

Using tangent words for curve segmentation

The length of the smallest word appearing in the decomposition of a smooth curve into maximal digital straight segments does not necessarilly goes to infinity when the mesh of the grid goes to zero (it is the case for strictly convex curves [ref]).

However it becomes the case if we replace digital straight segments by tangent words.

Using tangent words for curve segmentation



PhD proposal

A lot of open questions around the notion of tangent words are still open.

There is a grant for a PhD in Montpellier about this topic for the period 2011–2014. Please contact us if some good student could be interested.

`http://www.lirmm.fr/~monteil/proposition-de-these/`